MAY 23 1924

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS:

HAROLD ORDWAY RUGG, *Chairman.*
Lincoln School of Teachers College.
Teachers College, Columbia University.

JAMES CARLETON BELL,
Brooklyn Training School for Teachers.

FRANK NUGENT FREEMAN,
University of Chicago.

ARTHUR IRVING GATES,
Teachers College, Columbia University.

VIVIAN ALLEN CHARLES HENMON,
University of Wisconsin.

RUDOLF PINTNER,
Teachers College, Columbia University.

BEARDSLEY RUML,
Carnegie Corporation, New York City.

LEWIS MADISON TERMAN,
Leland Stanford University.

EDWARD LEE THORNDIKE,
Teachers College, Columbia University.

LAURA ZIRBES, *Assistant Editor.*
Lincoln School, of Teachers College.

VOL. XV.
No. 5

MAY, 1924

$4.00 a year.
60c. per copy.

CONTENTS

A New Measure of the Development of the Carpal Bones and Its Relation to Physical and Mental Development. *Frank N. Freeman and Thomas M. Carter* 257

Johnson's Measurement of Rate of Improvement Under Practice. *Joseph Peterson* 271

Language Handicap and the Stanford Revision of the Binet-Simon Tests. *Fred C. Walters* 276

Standardized Content in High School Biology. *R. M. Ruch and Leo H. Cossmann* 285

Mental Age and School Attainment of 1007 Retarded Children in Massachusetts. *Prepared under the direction of Amy Hewes by Mildred Holt, Sophie Meranski, and Julia Snell* 297

The Significance of Weighted Scores. *Paul V. West* . . . 302

Information as a Measure of Intelligence and Maturity. *William R. Wilson* 309

An Extensive Experiment in Motor Learning and Re-learning. *S. R. Braden* 313

Notes on Articles in Educational Psychology in Current Issues of Other Magazines 316

New Publications in Educational Psychology and Related Fields of Education 320

Published Monthly Except June to August by

WARWICK and YORK, Inc.,

York, Pa. *Baltimore, Md.*

...d as Second Class matter November 15, 1921, at the Post Office at York, Pennsylvania, under the Act of March 3, 1879.

The Journal of Educational Psychology

is published monthly, by Warwick & York, Inc., Baltimore, Md., and York, Pa. The issues for 1922 make Volume XIII. Title page and index are bound in December number.

Manuscripts for publication, books or other materials for review, and news items should be addressed to Harold Ordway Rugg, Chairman of the Editorial Board, 59 Edgecliffe Terrace, Yonkers, New York.

The price of the Journal is $4.00 a year, payable in advance. Foreign postage is 30 cents extra. Single issues are 60 cents each, and less than a full year's subscription is charged at single issue rate for every month ordered.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which change is to take effect. No claim for non-receipt of an issue will be entertained unless made within two weeks after receipt of the next succeeding issue.

Warwick & York. Inc.,
Baltimore, Md.

TEACHERS' HEADQUARTERS

European

Plan

Entirely

Fire-Proof

Centrally

Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

National Summer School

SPEND a few weeks in this delightful summering place and pursue credit courses under nationally famous teachers. "America's Greatest Summer Faculty" at your service. Special faculty includes: Allee, Chicago; Angell, Yale; Blackwelder, Stanford; Branson, N. Carolina; Carruth, Stanford; Cowles, Chicago; Franzen, California; Knappen, U. of Kansas; Merk, Harvard; McLain, Detroit; McCollum, Johns Hopkins; Reed, Ohio State U.; Shearer, Long Beach; Thorndike, Columbia; Turner, Harvard; Wood, Columbia. Special lecturers: Wm. G. Anderson, Liberty Hyde Bailey, David Starr Jordan, Shailer Matthews, E. A. Steiner, A. E. Winship and John Adams of London U.

Unique recreation features and excursions to world-famed scenic places.

(Special faculty and lecturers to serve through entire first 6 weeks.)

1st term: June 9 to July 18.
2nd term: July 21 to Aug. 29.

Register June 6 or 7. Fee $25 for 6 or 12 weeks.

Routing via Yellowstone Park, with stopover privilege at Logan, Utah, at no extra cost.

Vacation railway rates.

Write for Catalog.

Utah Agricultural College
Logan Utah

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XV | May, 1924 | Number 5

A NEW MEASURE OF THE DEVELOPMENT OF THE CARPAL BONES AND ITS RELATION TO PHYSICAL AND MENTAL DEVELOPMENT

FRANK N. FREEMAN

University of Chicago

AND

THOMAS M. CARTER

Albion College

The Problem

The investigation which is reported in this article is a study of physiological development and its relation to mental development.[1] We may introduce the report of the study by a comment on its relation to the previous investigation of this problem.

The study of physiological development has gone through two general stages. The first stage consisted of an investigation of the general course of development by plotting curves of age growth based upon the medians of successive ages. This kind of study is represented in the familiar growth curves in weight, height, vital capacity, arm span, and the like.

This type of measure of development meets the difficulty that one cannot readily interpret the stage of development of individuals by comparison with such a curve. If the height of an individual of a certain age is above the median height for that age we are unable to determine to what extent this superiority is due to precocity in development on the one hand, or to the fact that the individual will be taller than the average at maturity on the other hand. In other words, indi-

[1] Summary of part of a study reported by Thomas M. Carter in a doctor's dissertation which is on file in the library of the University of Chicago.

viduals differ in final size and in rate of development, and the mere comparison of individual measures with median growth curves does not enable us to distinguish these two factors.

Another way of putting the difficulty is to say that the composite growth curve which is made up from the medians includes within it a variety of types of growth curve and is to be regarded not as a representative of typical growth, but as a mere mathematical average of a number of types. As such, it is of little value as a means of individual diagnosis.

Because of the defects of the general growth curve which have been mentioned, it cannot readily be used as a means of determining the correlation between individual differences in the time or rate of physical maturing and individual differences in mental maturing or in intelligence. Some means must be found to distinguish between the two aspects of growth, namely: Final size and stage of maturity.

It may be remarked, parenthetically, that we meet a similar difficulty in our interpretation of measures of intelligence. It is customary to identify mental maturity with superiority in intellectual capacity. We have no means of distinguishing between the level of capacity which the individual will ultimately reach and the rate of maturing of intellectual capacity. This is a problem on which light may possibly be thrown by repeating measurements of children up to maturity.

To return to the discussion of physical measures, we come to the second stage in investigation. In this stage the attempt is made to determine individual rates of maturing independently of final size. The first method which was attempted was the study of individual growth curves and the identification of periods of acceleration which represent definite stages of maturing. Baldwin's survey of individual growth curves was the first wide scale attack upon this problem. This identification of stages of development requires possession of a series of repeated measurements on the same individual. There is some correlation between final size and the periods at which stages of development occur, and final size can be predicted fairly closely from the size at birth; but this correlation between size and rate of development is not close enough to be used as the basis of prediction.

It is therefore desirable to find other measures of physical maturity. This has led to the attempt to find in the ossification of the bones a maturity measure. The study by Rotch, Bean, Woodrow, and Baldwin, of the ossification of the carpal bones was for the purpose of dis-

covering in it the required measure. This study was at first pursued by the inspection method. The stages of development are distinguished by comparing the radiographs of the carpal bones of individuals with radiographs which are taken to represent the medians for the successive ages. The distinguishing marks of maturity are found in the appearance and the size of the bones. There is no means of determining which elements are most important as indices of maturity by this method, nor is it possible because of the lack of quantitative measurement to obtain a verification of the method.

An advance over the inspection method was made by Baldwin when he measured by means of a planimeter the area of projection of the carpal bones in the radiograph. This method meets one of the difficulties in that it makes the measure quantitative. It has another difficulty, however, in probably greater degree than the inspection method, on account of the fact that, like the earlier measues of height, weight, etc., it does not distinguish between size, and rate of maturing. That this is a serious difficulty will appear from the data of the present study which will be reported in a later paragraph.

Method of Present Study

This study is concerned, first, with a method of measurement by means of which the difficulties which have just been described might be overcome and with its verification by a comparison of the growth curves of boys and girls. It is concerned, in the second place, with the use of this method for the study of certain aspects of physical development and for the study of the relation between physical and mental development.

We are concerned first, then, with the method of measurement and with the examination of its validity. It was soon discovered from the study of the measurements of the areas of the projection of the carpal bones that this area was the product of two factors and that unless these factors were distinguished, this measure could not be interpreted. In order to distinguish these factors, it is necessary to secure an independent measure of general skeletal size. This was done by finding the area of a quadrilateral which was made by joining specified points on the bones at the extremities of the bones of the arm and of the hand. These parts will be specified more particularly in a moment. This area, as will be shown in a later article, was found to represent with considerable accuracy the general size of the individual. The assumption which was made as a starting point was that the size of the carpal

bones would be influenced by the general size of the skeleton and that this could be discounted by the use of the area of the above mentioned quadrilateral as a term in a ratio. . Specifically, the area of the carpal quadrilateral was used as the denominator of a fraction. The stage of development of the carpal bones was then represented by finding the ratio of the sum of the areas of the carpal bones to this carpal quadrilateral. To sum up then, we measured first the area of a quadrilateral made by drawing lines from fixed points in the wrist and calling this area the carpal quadrilateral. We then found the sum of the areas of the projections of the carpal bones and called this the total ossification. We then found the ratio of the total ossification to the carpal quadrilateral and called this the ossification ratio.

Technique of Measurement

It is readily recognized that there are three dimensions of the bones in question. For a complete knowledge of the amount of ossification present in any given unit the investigator should know all three of these dimensions. One is, however, limited if he uses only a single radiograph to two of these dimensions, namely length and breadth, since thickness cannot be skiagraphically determined. If the bones were regular in shape, the two dimensional diameters when considered together would afford as accurate a method of determining size of skiagraph as could be obtained. But the bones were found to be quite irregular. The diameters, therefore, might differ greatly while the areas of the two bones would be the same. We rejected, accordingly, the diameter method of determining the size of the bones.

It seemed evident that what was needed was some method which would measure the entire exposed area of the ossified surface. A planimeter was then secured and when its accuracy had been thoroughly tested experiments were made to determine its adaptability to our needs.

In order to test the accuracy with which this method of measuring could be used in our investigation, one of the authors selected 50 skiagraphs and measured them twice and calculated the coefficient of correlation between the two series of measures. That the manipulator might not be influenced in the second measurement by the results obtained in the first measurement, the entire series of bones was measured for the first time and then for the second time. To guard further against bias which might result from knowing what measurements were obtained, the planimeter readings were dictated to another person

and the exact measures deduced from the readings were not known by the manipulator until the two series of planimeter tracings had been completed. The coefficient of correlation obtained was .987 ± .009. The results obtained were as accurate as could be expected and the method was adopted as satisfactory for determining the size of the skiagraphs of the bones.

The skiagraphic area of the bones was measured from the films rather than from the prints. For illuminating the films from beneath, a box, 11 inches wide, 10 inches deep and 22 inches long over all, was constructed and provided with electric fixtures for two lamps. The box was enameled white inside and the top was made of frosted glass. Such an arrangement provided ample illumination, but the heat produced by the two lamps was greater than the films could stand. Accordingly a piece of plate glass the size of the top of the box and raised from the frosted glass top by supports at each end about an inch high was added to the apparatus. A sufficient air chamber was thus provided between the box and the plate glass so that the heat was no longer a disturbing element. Not only did this device solve the problem of heat, but it also provided an easy arrangement for anchoring the films and holding them steadfast while the measurements were being made. The films were fastened to the piece of glass by strong spring clamps.

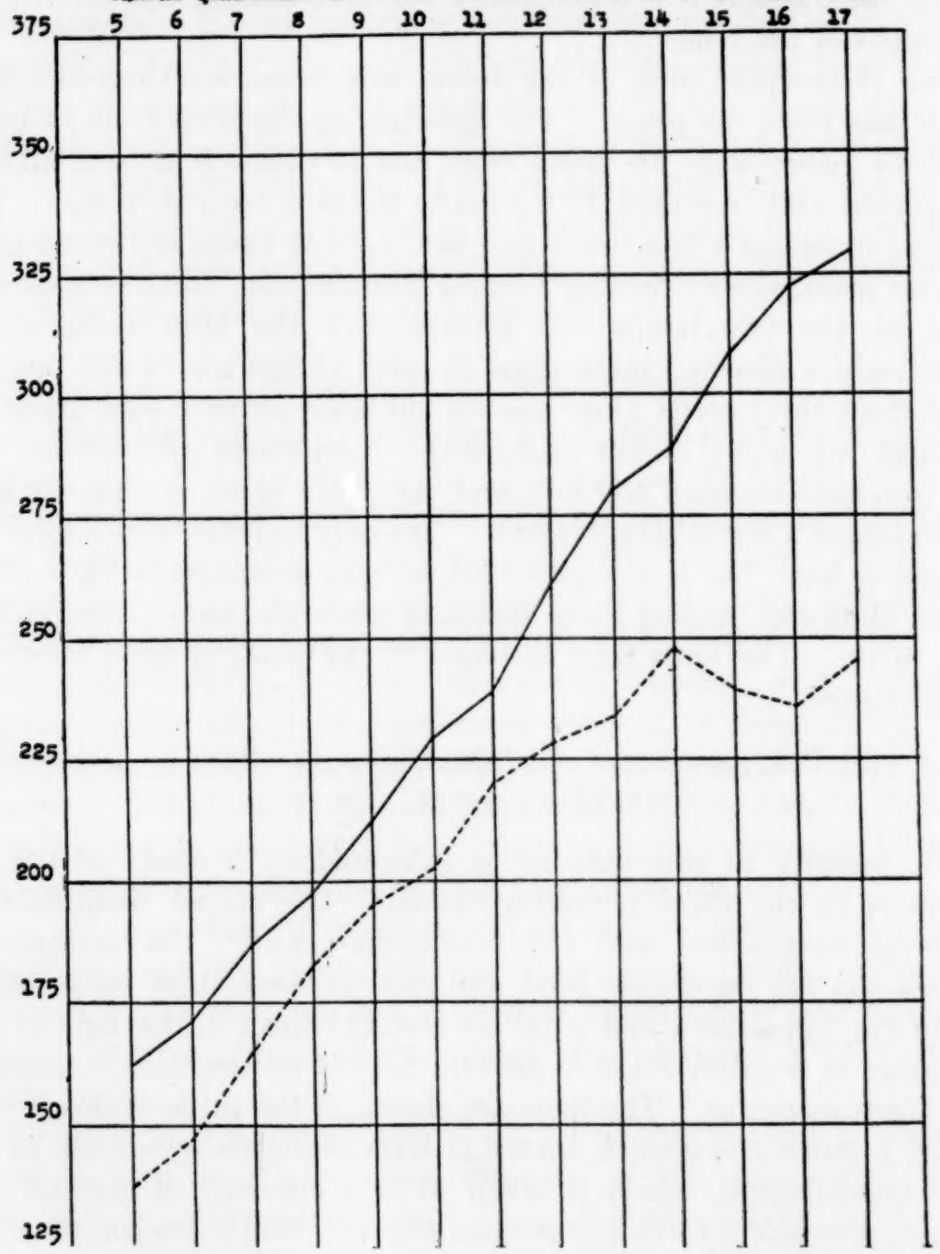
The Validity of the Ossification Ratio as a Measure of Maturity

The validity of this measure is indicated by a study of the sex differences in the three measures, namely: The carpal quadrilateral, the total ossification, and the ossification ratio. The comparison between the age curves for boys and girls in these three measures fit exactly our hypotheses and what we should expect in the light of our knowledge of sex differences in growth and our assumptions concerning these three measures. The facts are shown in the three charts.

Fig. 1 shows the growth curves of boys and girls separately in the carpal quadrilateral, which is taken to be a measure of skeletal size. The characteristic growth phenomena are here well represented. The boys' measures are larger than those of girls, as we find in the case of height and weight. The acceleration in case of girls comes earlier than in the case of boys. The number of cases for each age (20 for each sex) is not large enough to give a perfectly smooth curve, but it will be seen that the curve for girls approaches that for boys in the pre-adolescent

period but that the boys' curve diverges from that of girls beginning with 11 years. At this point the rate of growth for girls begins to diminish.

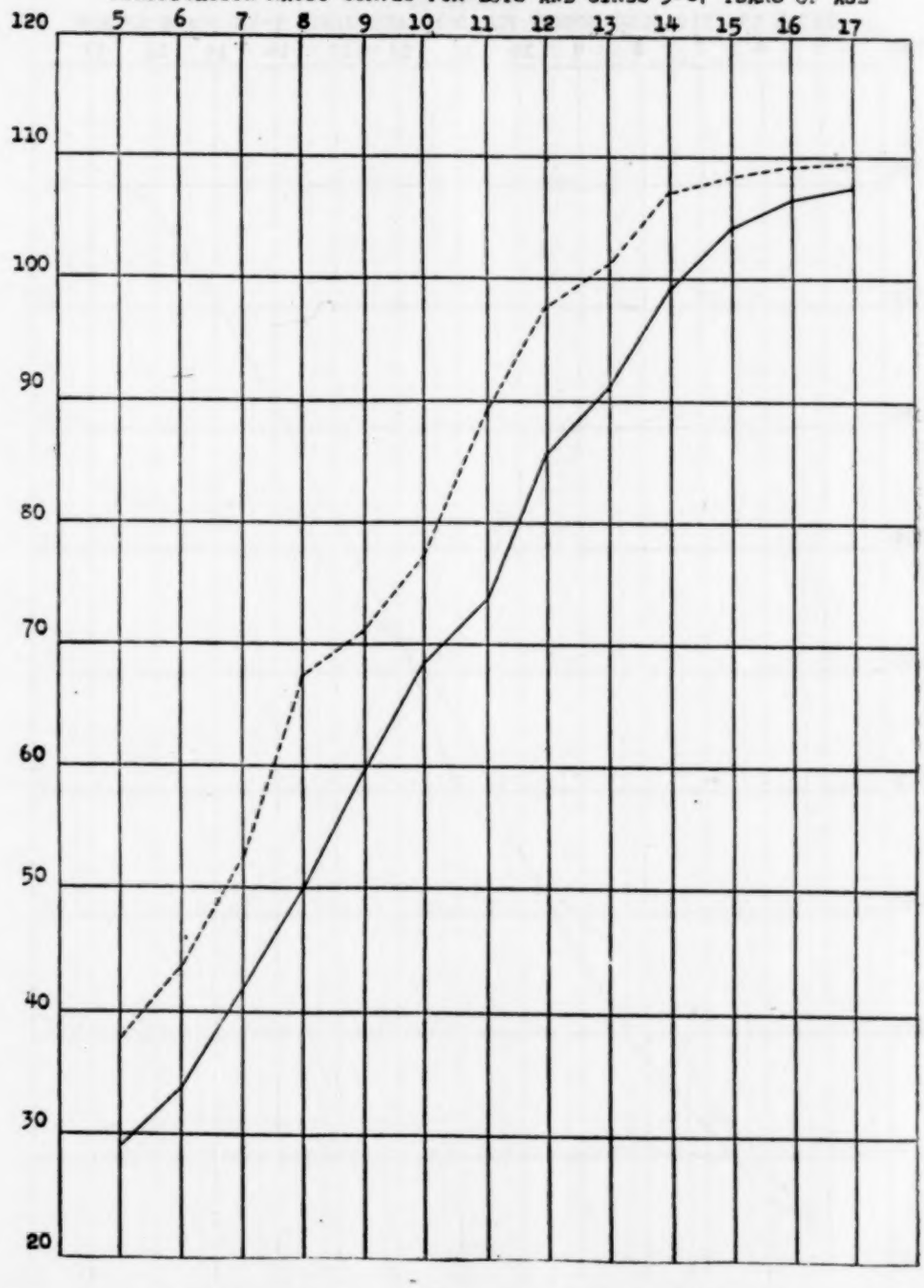
FIGURE 1.
CARPAL QUADRILATERAL CURVES FOR BOYS AND GIRLS 5-17 YEARS OF AGE



Contrast these curves with the curves for ossification ratio of boys and girls shown in Fig. 2 Here the curve for girls is above that for boys at every age, as we should expect. Furthermore, the pubertal acceleration comes a year earlier for girls than for boys, namely at 10

for girls and 11 for boys. The development of boys, however, continues longer than that for girls and the appearance of the curves suggests that they would come together at some later age, probably 19

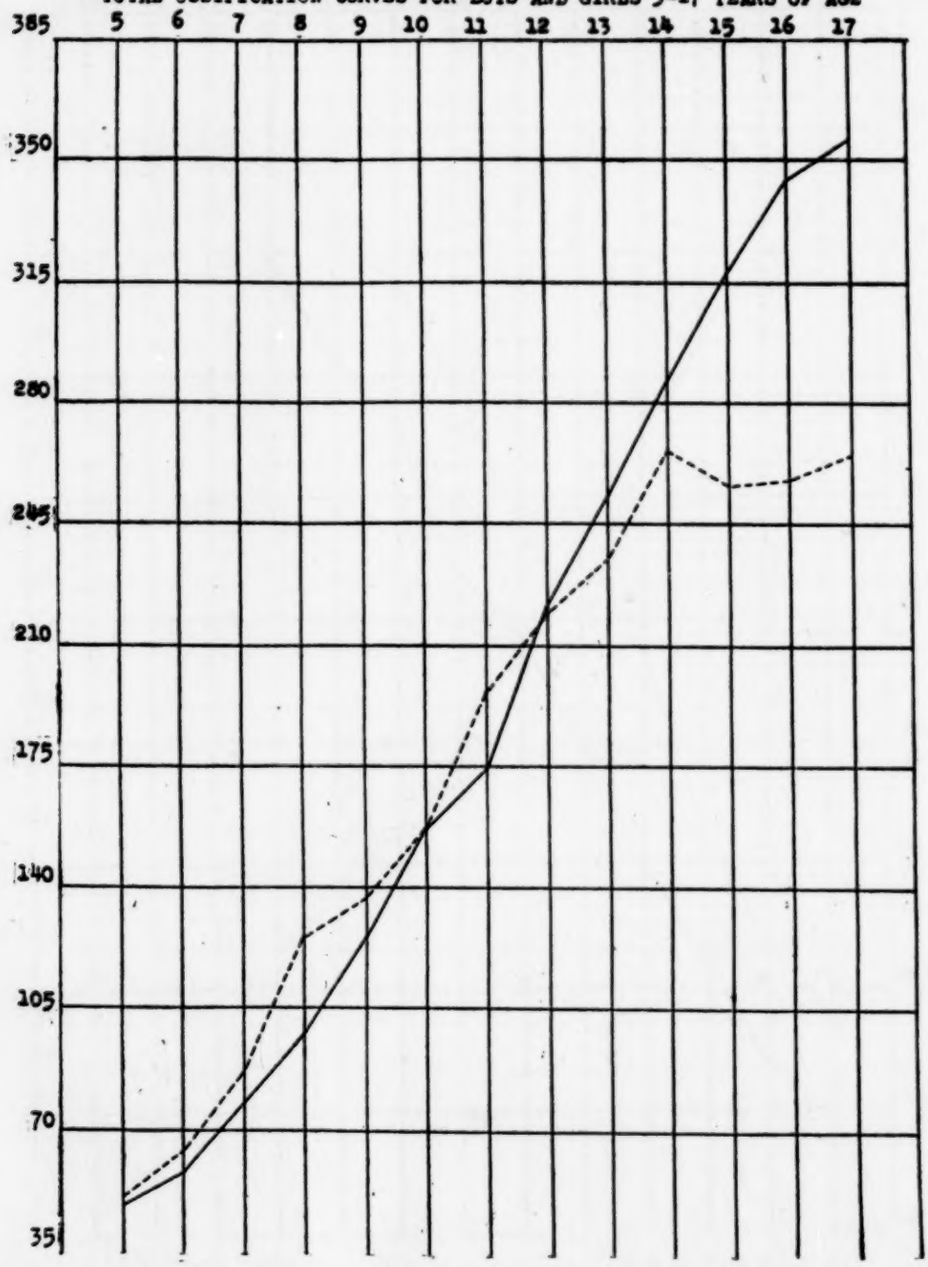
FIGURE 2.
OSSIFICATION RATIO CURVES FOR BOYS AND GIRLS 5-17 YEARS OF AGE



or 20. The girls show a rate of development which is from one to two years in advance of that of the boys. All this agrees exactly with our general evidence concerning the relation between the rate of maturing of the two sexes.

Contrast the curves of the ossification ratio with the curves of total ossification shown in Fig. 3. The curves for boys and girls cross and re-cross until age 14. The irregularities in the curve are probably due in part to the limited number of cases, but the general fact up to

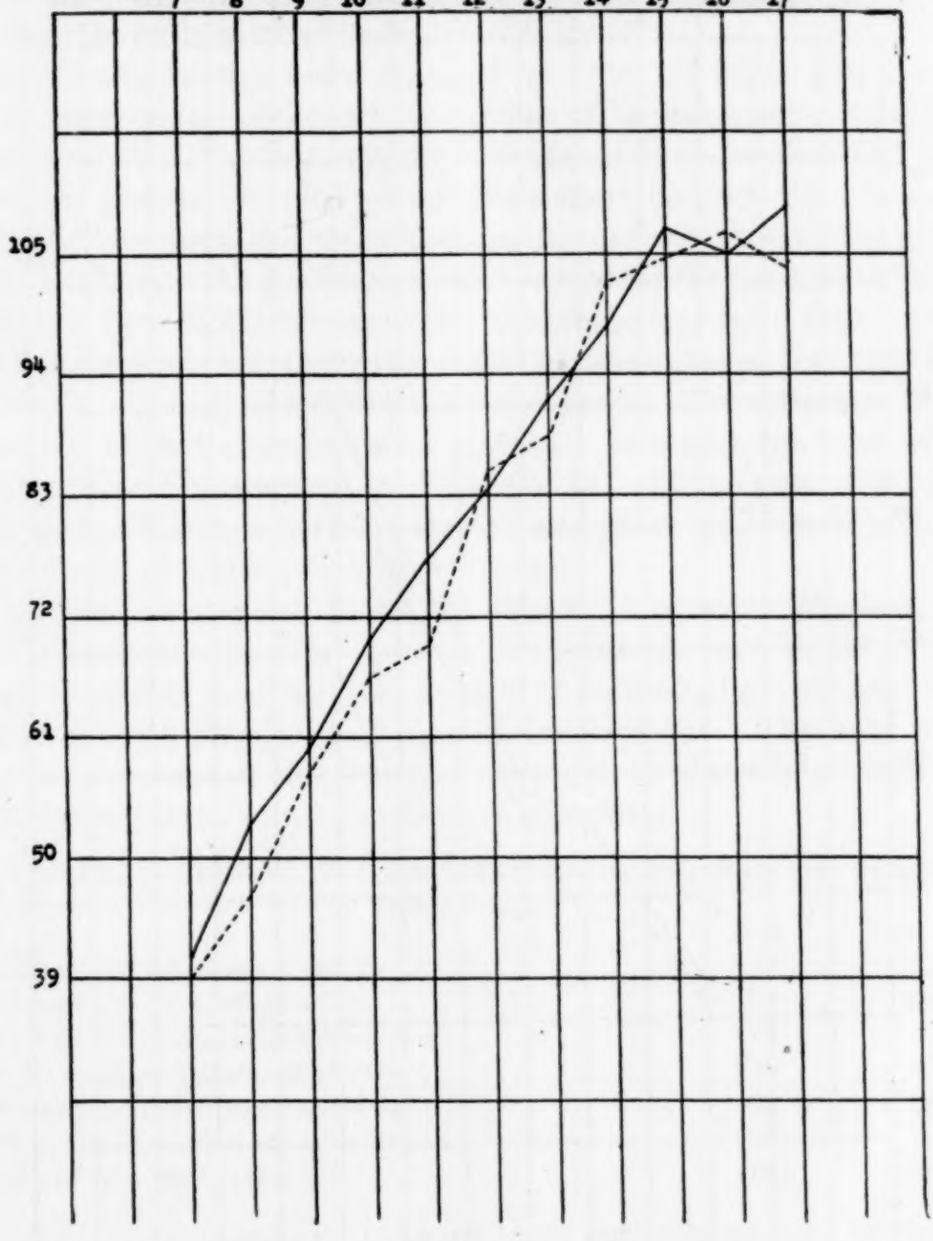
FIGURE 3.
TOTAL OSSIFICATION CURVES FOR BOYS AND GIRLS 5-17 YEARS OF AGE



adolescence is that, so far as this measure is an indication, the development of boys and girls is practically alike. This we know to be contrary to the facts and therefore a total ossification is not a suitable measure so far as the comparison between boys and girls is concerned.

If the ossification ratio is valid as a comparative measure of the development of boys and girls by virtue of the fact that it enables us to discount differences in size and to obtain a measure of development

FIGURE 4.
OSSIFICATION RATIO CURVES FOR TALL AND SHORT BOYS 7-17 YEARS OF AGE

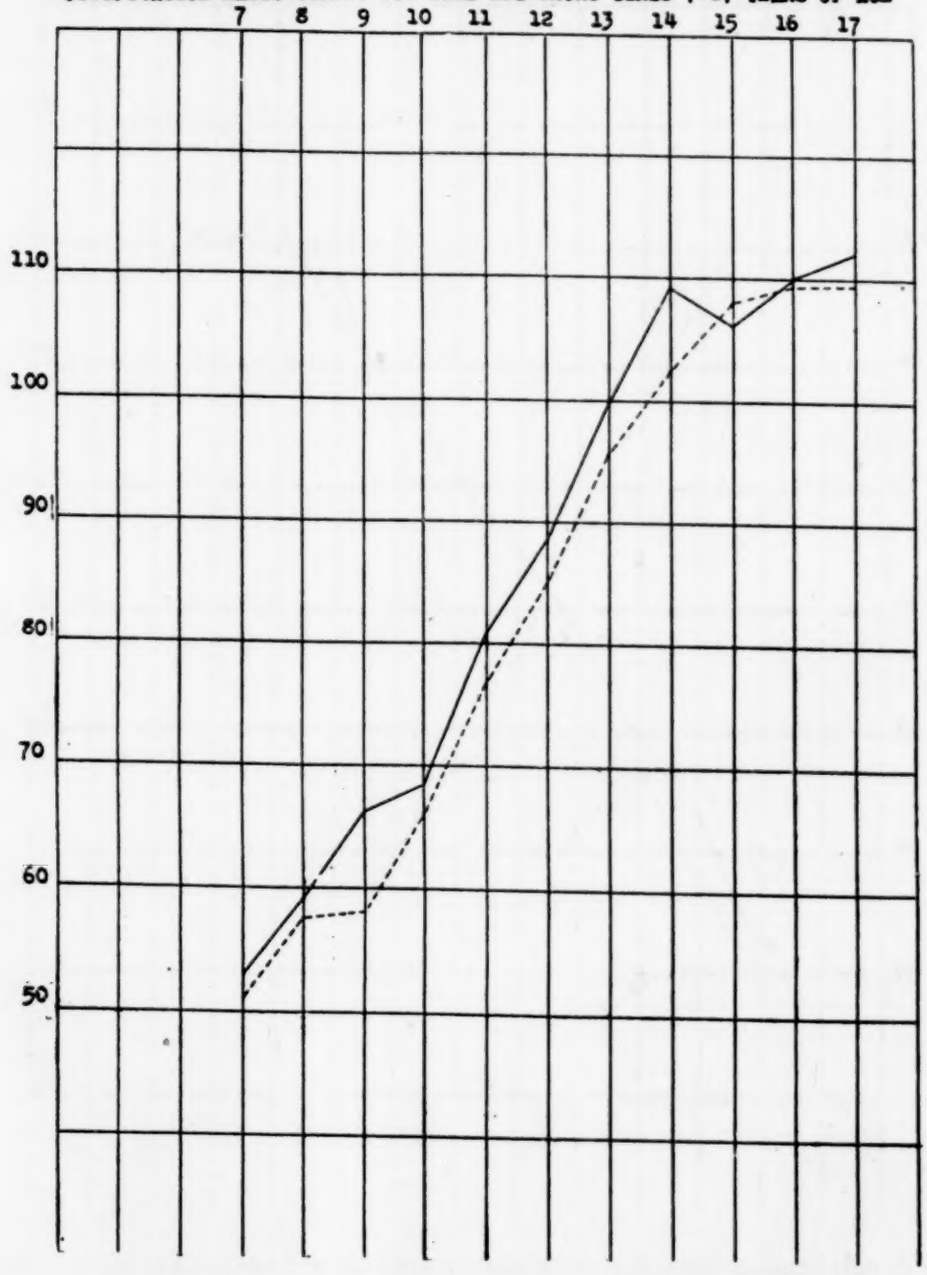


unaffected by size, we may infer that it will also serve as a satisfactory means of measuring the relative maturity of individuals of the same sex regardless of their differences in size.

Preliminary Results from the Use of the Ossification Ratio as a Measure of Maturity

Some of the results of the use of this measure with respect to physiological development have already been mentioned in commenting on

FIGURE 5.
OSSIFICATION RATIO CURVES FOR TALL AND SHORT GIRLS 7-17 YEARS OF AGE



the curves of the ossification ratio. The stages of development appear to be comparable to those which we find when we measure the acceleration in height. Apparently the acceleration in development of the

carpal bones runs parallel to the acceleration in general physical growth. The most interesting feature of these curves as measures of development is that they give an appearance of ultimately coming together. If it is a true measure of development we should expect this union of the curves to occur. It is further interesting to note that the point at which development ceases and at which the curves will unite is apparently several years beyond age 17.

The measurements were classified in order to bring out a comparison between tall and short individuals of the same sex. It will be remembered that Baldwin has given evidence to show that tall individuals, in general, develop earlier than short individuals. In order to get evidence upon this point the age curves for each sex were classified by height and the ossification ratio curves for the tall and the short individuals were plotted separately. As is apparent from Figs. 4 and 5, the tall individuals are somewhat earlier in development than the short individuals, although the difference is not great. The difference in the case of boys is not apparent after year 11. It is more uniform in the case of girls and appears up to the age of 14. In the case of the girls the average difference between the tall and short individuals is in the neighborhood of a half year's development.

The last comparison to be mentioned is between physical and mental development. The data on this point are provisional and are subject to further analysis from a study of individual growth curves on the side of both physical and mental development. The data which we have to present at the present time are in the form of correlations. The following table forms the basis of discussion.

CORRELATIONS BETWEEN VARIOUS MEASURES OF PHYSICAL AND MENTAL DEVELOPMENT

	BOYS	GIRLS
Mental age and ossification ratio	.73	.75
Mental age and chronological age	.82	.83
Chronological age and ossification ratio	.85	.87
Grade location and anatomical ratio	.86	.89
Grade location and chronological age	.93	.94
Partial correlation with age constant		
Mental age and ossification ratio	.084	.088

One or two comparisons between these correlations may be pointed out. Chronological age seems to have a closer correlation with mental age than does the ossification ratio. This would seem at first sight to contradict the opinion that mental development corresponded more closely to physiological maturity than it does to age. At any rate, it

seems to make it improbable that we may find a measure of physiological maturity which will give a better base upon which to calculate intelligence than is furnished by chronological age. It may possibly be that the closer correlation between chronological age and mental age is due to the fact that mental age has not been analyzed into the two factors of degree of maturity and final intelligence, though it is not entirely clear why this should favor the correlation with chronological age rather than with anatomical ratio. This is a question which demands further investigation.

The closer relation between grade location and chronological age than between grade location and anatomical ratio may be due to the fact that chronological age is used as a basis for entrance to school and for promotion. In both cases the correlation is high.

When we partial out age and find the correlation between mental age and ossification ratio with chronological age constant, we get practically a zero correlation. This coefficient agrees closely with that found by Baldwin and Stecher, (p. 57) which was .09.[1] This means that those individuals who are precocious in physiological development are not the individuals who are high in intelligence in comparison with others of the same age.[2]

Again we have the difficulty in interpretation due to the fact that the measure of mental age is a composite. Ossification ratio is a measure of the distance which the individual has traveled along the road to maturity, or precocity. Mental age, on the other hand, is a measure both of the distance the individual has traveled and of the level of the road along which he is traveling—of precocity *and* brightness. It is a composite measure of the two and we do not know in any case how much it may be affected by each of these two factors. In such a case our interpretation of the correlation between ossification ratio or anatomical maturity and mental age is somewhat involved, but the following inference seems to be sound.

---

[1] Baldwin, B. T. and Stecher, Lorle I.: "Mental Growth Curve of Normal and Superior Children, Iowa City." University of Iowa Studies, Vol. II, No. 1, 1922.

[2] Since the above was written, the writers have received the monographs of Prescott in which very low partial correlations between anatomical index and mental age with chronological age constant are reported. For four groups the coefficients are, respectively, −.05, .12, .30, .33. See Prescott, D. A.: "The Determination of Anatomical Age in School Children and Its Relation to Mental Development." Cambridge, Harvard Monographs in Education, 1923, Series 1, No. 5.

Suppose that anatomical precocity is correlated with mental precocity. Anatomical precocity should then be correlated with mental age. We have seen that mental age is made up of the two factors of brightness and precocity. Anatomical precocity may or may not be positively correlated with brightness. It is almost certainly not negatively correlated with it. If, then, anatomical precocity is correlated with one of the two factors in mental age—*viz.* mental precocity, and if its correlation with the other is at least not lower than zero, it should be positively correlated with the composite of the two. Since anatomical precocity is not correlated with the composite (mental age) it cannot be correlated with mental precocity, which is one of the two factors in this composite.

The same argument applies to the other factor of mental age—brightness. Anatomical precocity appears not to be correlated with it either.

These conclusions seem at first sight to be contrary to certain fairly well established facts. A number of studies, culminating on the recent extensive survey by Terman, indicate that bright and scholastically advanced children are taller and heavier than are dull and retarded children. Upon analysis, however, it will be seen that fact is not necessarily contradictory to our conclusions. Size and weight themselves are the product of two factors, final size and rate of growth. Only rate of growth corresponds to anatomical precocity. The facts do not prove that rate of physical growth is correlated with the rate of mental growth. We have been inclined to interpret our facts as indicating such a correspondence, but the lack of correlation between ossification ratio and mental age at least casts grave doubt upon this interpretation.

A word as to the practical bearing of these theoretical conclusions. Our data seem to indicate that individual variations in the rate of skeletal development are independent of individual differences in the rate of mental development. This fact bears, first, upon our method of reckoning intelligence or brightness. Our method has been to reckon intelligence with reference to chronological age as a base—as by finding the ratio of mental age to chronological age. It has been proposed that we substitute some measure of physical development—as anatomical age—for chronological age, in this calculation. It seems quite clear from our data that we should not be justified in making such a substitution. There is an error in our present calculations, but there would be a larger error in the proposed calculation, because mental

age is more closely correlated with chronological age than it is with ossification ratio.

Again, it has been suggested that children be classified in school on the basis of their physical development, as measured by such an index as ossification ratio. Two issues are involved, the adjustment of school organization to social development, on the one hand, and to intellectual development, on the other hand. Promotion according to anatomical development is proposed on the ground of its relationship both to social and to intellectual development. We are here concerned only with the latter. Our data show that the child's capacity to do advanced intellectual work cannot be measured by his rate of anatomical development. The child's variation in mental development must be measured by mental tests, and it cannot be inferred from the results of physical tests. Classification should be based primarily on measures of intellectual growth and capacity, and only secondarily, if at all, upon measures of physical growth and capacity.

JOHNSON'S MEASUREMENT OF RATE OF IMPROVEMENT UNDER PRACTICE

JOSEPH PETERSON

George Peabody College

An important though not uncommon error lurks in the data and arguments which Johnson presents, in his very interesting papers.[1] His view is that brighter subjects improve faster than dull ones, a view

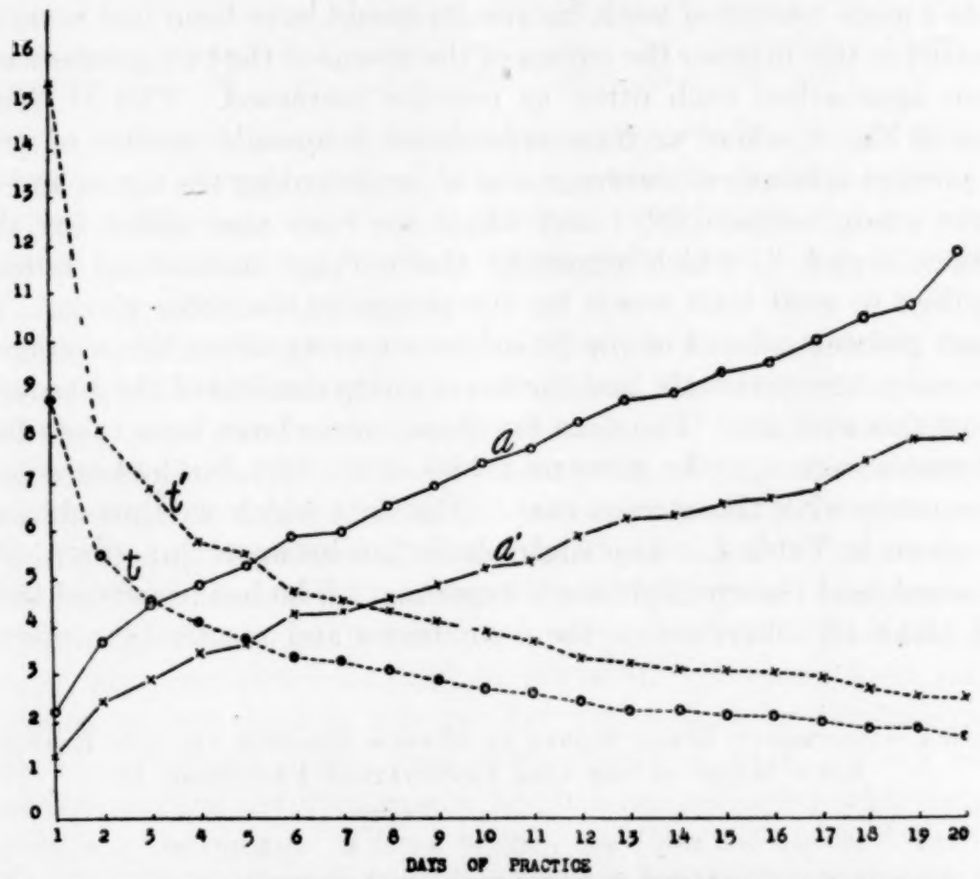


FIG. 1.—*A* and *A'* show improvement in learning of the 30 subjects who had scored above the group average in the intelligence tests and of the 30 who had scored below average, respectively. Units give the average number of words per day's practice of 10 minutes, in terms of hundreds. Curves *t* and *t'* show for these groups in the same order the time required to read 1000 words, units being in 5-minute periods.

that may or may not be true, however, according to how special the field of training is. Positive correlations between intelligence test scores and improvement in ability to read "were obtained by measuring

[1] Johnson, O. J.: A Study of the Relation between Ability to Learn and Intelligence as Measured by Tests. *Journal Educational Psychology*, Vol. XIV, 1923, pp. 540–544.

the improvement in terms of the average number of words read during the first three days and the last three days of practice" (p. 451). As he unfortunately does not publish the individual scores we can not make the specific corrections desirable in his correlations. His curves show that the difference between the mean scores of the brighter and the duller groups, as determined by the tests, increases with practice. This he takes as evidence that the brighter subjects improve more rapidly; but it is of course obvious that if instead of using *the amount of work done in* 10 *minutes* as his measure he had taken *the time required to do a given amount of work*, his results would have been just reversed. Plotted in this manner the curves of the means of the two groups would have approached each other as practice increased. This is clearly seen in Fig. 1, where we have reproduced Johnson's amount of work, or average attainment curves, a and a', representing the upper and the lower group, respectively; and where we have also added the time curves, t and t', which represent the average number of minutes required to read 1000 words by the groups in the order given. The upper group consisted of the 30 subjects scoring above the average of the entire 60 individuals, and the lower group consists of the 30 scoring below this average. The data for these curves have been taken from Johnson's curves, as he gives no tables of his data, but measurements were made with the utmost care. The data which we thus obtained are given in Table I. Any slight deviation between these data, thus obtained, and those of Johnson's experiment if he has preserved them, will make no difference in the conclusions and results here offered.

TABLE I.—SUCCESSIVE MEAN SCORES IN MIRROR READING AND THE RATIOS OF EACH MEAN TO THE ONE IMMEDIATELY PRECEDING IT

	Trials									
	1	2	3	4	5	6	7	8	9	10
Average number of words per 10 minutes										
Upper group.....	223	365	442	484	527	587	597	652	690	740
Lower group.....	130	246	292	346	364	424	437	460	488	522
Average number of minutes to read 1000 words										
Upper group.....	44.8	27.4	22.6	20.7	19.0	17.0	16.7	15.5	14.5	13.5
Lower group.....	76.9	40.7	34.2	28.9	27.5	23.6	22.9	21.7	20.5	19.2
Ratios of "average number of words" scores										
Upper group.....	....	1.64	1.21	1.10	1.09	1.11	1.02	1.09	1.06	1.07
Lower group.....	....	1.89	1.19	1.18	1.05	1.16	1.06	1.05	1.06	1.07
Ratios of "average number of minutes" scores										
Upper group.....	....	.61	.82	.91	.92	.90	.98	.92	.94	.93
Lower group.....	....	.53	.84	.85	.95	.86	.94	.95	.94	.93

	Trials									
	11	12	13	14	15	16	17	18	19	20
	Average number of words per 10 minutes									
Upper group.....	772	824	874	883	925	944	998	1042	1068	1184
Lower group.....	540	593	622	637	657	668	685	747	785	797
	Average number of minutes to read 1000 words									
Upper group.....	13.0	12.1	11.4	11.3	10.8	10.6	10.0	9.6	9.4	8.4
Lower group.....	18.5	16.9	16.1	15.7	15.2	15.0	14.6	13.4	12.7	12.5
	Ratios of "average number of words" scores									
Upper group.....	1.04	1.07	1.06	1.01	1.05	1.04	1.06	1.04	1.02	1.11
Lower group.....	1.03	1.10	1.05	1.02	1.03	1.02	1.02	1.09	1.05	1.01
	Ratios of "average number of minutes" scores									
Upper group.....	.96	.93	.94	.99	.95	.96	.94	.96	.98	.90
Lower group.....	.97	.91	.95	.98	.97	.98	.98	.92	.95	.99

It is clear that the time curves taken on their face value, as Johnson has taken the average attainment curves, would indicate just the opposite conclusion from that which he drew. Which kind of data then shall we use? Obviously neither, in the light of this contradiction, if we are to be safe in our conclusions.

Taking Johnson's own curves, a and a', we should certainly find it much fairer to determine the superiority of the better group average over the poorer one in terms of percentage, and then to compare the final with the initial difference in this respect. By this method we find that the superior group starts with a mean number of words per 10 minutes that is 49 per cent greater than the mean of the poorer group, when the first three trials are considered but they have a mean superiority in the average of the last three trial scores—the eighteenth, nineteenth, and twentieth—of only 41 per cent. It would seem fair to measure improvement in this manner, rather than in absolute gains, since improvement means gain over something as a starting point: And this method involves us into no such inconsistency with the time curves, or the means of time scores, as does the method used by Johnson. These figures, then, contradict Johnson's conclusion that the upper group improved more than the lower group. So also we find that the increase in the upper group curve is 415 per cent while that of the lower group, curve a', is 514, the increase being the gain over the point of beginning, or over the first mean score, in the case of each group. If, to get more stable measures for comparison, we take the average of the first three and of the last three trials, we get respective increases of the upper and the lower groups of 209 and 249 per cent. This method also shows greater improvement of the lower group. Moreover, the harmonic mean of the increase of each trial, from the

second to the twentieth, over its preceding trial in terms of mean number of words per 10 minutes—that is, the harmonic mean of 19 increments—is 32.8 words for the upper group and 23.6 for the lower group. These means are increases over the respective starting average scores of 14.7 per cent for the upper group and of 18.2 per cent for the lower group of subjects.[1] All these data show that the lower group, the 30 subjects scoring below average in the intelligence tests, made more improvement in the learning than did the upper group of 30 subjects.

The geometric mean of the percentage increments or decrements shown in Table I, that is the nineteenth root of the product of the 19 percentages, would seem to be a fair measure of the improvement, in as much as it is the mean percentage increase of each stage of learning over the preceding stage. In average number of words per 10 minutes the geometric mean of the improvement of the upper group is 9.18 per cent as against a mean improvement of 10.01 per cent for the lower group. The corresponding geometric means of the decrease in average number of minutes to read 1000 words are 8.4 per cent and 9.1 per cent. These means also show greater gain for the lower group, contrary to Johnson's conclusions.

It is instructive to examine the actual ratios from which these geometric means have been found. These ratios are given in Table I. These data show more rapid improvement by the lower than by the upper group in the early part of the learning. Another way[2] to bring out this same fact of more rapid initial gain by the group scoring below average in the intelligence test is to determine which group reaches first the geometric mid-point of its own entire improvement, or of the improvement indicated by the final mean score as compared with the initial mean score. This mid-point is obtained by taking the square root of the product of the first and last mean of each group. We thus obtain a mid-point of improvement score for the upper group of 513.84 words in 10 minutes, and for the lower group of 321.88 words in 10 minutes for the lower group. Now, while there are technical objections to taking these as the actual mid-points of improvement—a

[1] The percentage of the lower group passing the median or the mean of the upper group would be a good measure to use, but we have not the necessary data for this method. If the upper group is learning faster than the lower one, a smaller percentage of the latter group would pass the average of the upper group toward the end than at the beginning of the experiment.

[2] For this suggestion and for other aid in the preparation of this paper I am indebted to my son, J. Kimbark Peterson.

position that would commit us to the view that the learning actually goes on much more rapidly in the early part of the process than in the latter part—the scores thus obtained at least represent comparable stages in the process. An examination of Table I shows that the lower group reached their "mid-point," say 322, at almost exactly the beginning of the fourth day of practice, whereas the upper group reached the corresponding point in their progress, say 514, at about .2 of the way beyond the beginning of the fifth practice period. Thus the lower group shows signs of having been less practiced at the beginning of the experiment on the mental function or functions to be trained in it, than was the upper group. Is it not possible also that this function or these functions are also represented in the intelligence tests to the extent that the difference in the scores of these two groups may have been in large part due to the greater practice of the so-called superior group. This is only a speculation, but it bears some support in the fact of greater initial improvement of the lower group. It raises a very interesting and important problem in intelligence testing and suggests a method of approach to its solution. We are undoubtedly taking too little account in tests of the respective amount of practice in the functions involved in them by the different subjects.

It is not surprising that the subjects scoring lower in the initial period of the learning should show greater improvement than the group scoring higher. It is known that practice of several subjects in a given function, or at least in some functions, makes them become more and more alike in respect to this function,[1] though it is a frequent fallacy on the part of experimenters to misinterpret the data by measuring results in absolute gains in terms of amount of work done in a given time unit, without considering the other facts to which attention has been called in our present discussion.

[1] Peterson, Jos.: Thurstone's Measures of Variability in Learning. *Psychological Bulletin*, vol. XV, 1918, pp. 452–455, where other references are given.

LANGUAGE HANDICAP AND THE STANFORD REVISION OF THE BINET-SIMON TESTS

FRED C. WALTERS

Graduate Student, Teachers College, Columbia University

The writer gave the Stanford Revision of the Binet-Simon Tests to 165 children in one of the New York City schools. Ninety-eight of these children were in Grade VIB and sixty-seven were in VIIA. The community in which the school was located was made up largely of foreign-language-speaking homes. It was thought that this foreign language factor would have some influence on the scores made on the tests by these children; so certain preliminary data were secured as the tests were given for the purposes of estimating this influence. Each child tested was asked the following questions: What language is spoken in the home? Is any other language spoken by either of the parents? If so, what language? To what extent is it used when speaking to you? Do you speak any other language than English? Do you understand the foreign language your parent or parents speak? Have you ever attended any other school than the public school? Where there was an affirmative answer to this question, it was followed up by other questions to find out what school, the extent of attendance, and the extent, if at all, to which this attendance interfered with attendance at the regular public school. Fourteen children reported Saturday attendance at a Bohemian-language school.

All of the children tested with the exception of two—one born in England and one in Canada—were born in this country; all but two or three had gone to an American public school for the term of their school-life; and all spoke fair unbroken English. The nationalities represented were Bohemian 47, German 43, Italian 16, Polish 5, French 1, Norwegian 1, Jewish 1, and American 51 (American here is used to designate all those of mixed extraction where English only was spoken by the parents). The occupations represented by the fathers of these children were: mechanic 14, carpenter 12, painter 8, metal worker 8, button factory worker 6, grocer 5, printer 5, marble worker 5, butcher 5, baker 5, electrician 5, automobile driver 8, salesman 5, brick layer 4, construction foreman 4, tailor 4, clerical worker 4, shoe factory worker 4, members of fire department 3, policeman 3, motorman 2, expressman 2, plasterer 3, common laborer 14, and one each of a dozen other trades or occupations.

As a means for determining the influence of foreign language usage on the scores made on the tests, the test booklets were divided into two groups on the basis of the language predominantly spoken in the home. That is, if the home language was quite generally English, the booklets were put into one pile; if, however, a considerable amount of the home language was a foreign language, the booklets were put into a second pile. These two piles or groupings will hereafter be spoken of as Group A (the English-speaking-home group) and Group B (the foreign-language-speaking-home group). For the purposes of still further refinement of the study there was rejected from Group A all booklets where any language other than English was spoken by the parents, whether this language was reported to have been used in the home or not. And there was rejected from Group B all booklets where the home language was not predominantly foreign. These last two groups will hereafter be spoken of as Group C and Group D, respectively. Groups A and B include all of the 165 children tested and represent a rough division of these children into those from English-speaking homes and those from foreign-language-speaking homes. Groups C and D are those children of groups A and B, respectively, where there was no question about the language spoken in the home being either all English or very largely foreign. Groups C and D should be thought of as refinements of groups A and B.

The Composition of the Groups

There was no marked difference on the basis of occupations of the fathers of these children between the groups. The trades and skilled occupations were fairly proportionately distributed between groups A and B, and C and D. The three members of the city fire department, the three policemen, and the two motormen all came from Group A; while the grocers, butchers and carpenters were found largely in Group B. The number of skilled occupations represented by Group A was 25; the number represented by Group B was 27. If a trend of occupations might be said to exist for either of the groups, it was in the direction of the clerical and semi-official for Group A and in the direction of the trades for Group B. The number in each group, the range of chronological age for each group, the mean chronological age, and the mean mental age and IQ as shown by the tests are given in Table I. The last two rows of Table I, headed "Group A – B" and "Group C – D," are differences and should be read "Group A minus B." and "Group C minus D."

TABLE I

Groups	Number in group	Range of chronological ages	Mean CA	Mean MA	Mean IQ
Total group	165	10-4 to 15-5	12-8.4	12- 2.3	96.3
Group A	67	10-4 to 15-5	12-6.6	12- 9.8	102.8
Group B	98	10-8 to 15-3	12-9.6	11- 9.2	91.9
Group C	51	10-4 to 15-5	12-7.3	12- 8.4	101.7
Group D	76	10-8 to 15-3	12-9.5	11-10	92.3
Group A − B	...	.....	0-(−3)	1- 0.6	11.0
Group C − D	...	.....	0-(−2.2)	0-10.4	9.4

PROCEDURE FOLLOWED

The general plan of procedure was the comparison of scores made on what might be termed non-language tests with scores made on tests presupposing average ability in English usage. The assumption underlying this comparison was the presupposition that what the children did on the non-language group of tests might be taken as a basis for estimating what they would do on the total group of tests if there were no language handicap operating. The group of tests that will hereafter be spoken of as the Non-language group and which was taken as the basis of comparison was composed of the Designs Test of Year X, the Ball and Field, Repeat Five Digits Backwards, and Pictures (interpretation) Tests of Year XII, the Induction Test, Arithmetic Reasoning, and Clock Tests of Year XIV, and the Problem of the Enclosed Boxes, Repeat Six Digits Backwards, and Code Tests of Year XVI. The tests selected because of the understanding and use of English which they presuppose were the Vocabulary and Naming 60 Words Tests of Year X, Vocabulary, Abstract Words, and Fables Tests of Year XII, Vocabulary Test of Year XIV, and Vocabulary, Interpretation of Fables, and Difference between Abstract Words Tests of Year XVI. These tests will hereafter be spoken of as the Language group.

All of the 165 children were given the tests for years X, XII, XIV, and XVI, excepting such children as passed all of the tests for Year XII, which year was usually given before giving the tests for Year X. Where all of the tests for Year XII were passed, full credit was given for Year X. Each test of each of the four years was scored as "passed"

or "failed," and these scores were tabulated by years. The percentage of children passing each of the tests for each of the four years was then determined. To make this clear, take Test I of Year XII. One hundred eighteen of the 165 children passed this test. Thus, the percentage of "passed" was 71.5. The percentage passing Test 2 of Year XII was 65.4, etc. These percentages are given in Tables II, III, IV, and V. The last two rows of these tables are differences and show the percentages of children in one group passing the particular tests in excess of the percentages of children in the contrasted group passing the same tests. For example, in the row headed "Group A – B,"

TABLE II

Year X

Name of test	Percentages of children passing the tests					
	Vocabulary	Absurdities	Designs	Reading and Report	Comprehension	60 Words
Total group	100	96.4	87.8	97	80	78.8
Group A	100	97	91	100	92.5	91
Group B	100	96	86	95	71.4	70.4
Group C	100	96	88	100	92.1	91.1
Group D	100	95	87	93.4	74	72.5
Group A – B	0	1	5	5	21.1	20.6
Group C – D	0	1	1	6.6	18.1	18.6

TABLE III

Year XII

Name of test	Percentages of children passing the tests							
	Vocabulary	Abstract Words	Ball and Field	Dissected Sentences	Fables	Five Digits Backwards	Pictures Interpretation	Similarities
Total group	71.5	65.4	58	56.5	55.8	70	60	55.2
Group A	91	83.6	68.2	71.6	71.6	64.2	64.2	74.6
Group B	56.1	51	53	44	42	71.4	54	42
Group C	90.2	82.4	58	70.6	70.6	60.8	60.8	72.6
Group D	55.2	47.4	51.3	42	27.6	72.3	52.6	39.4
Group A – B	34.9	32.6	15.2	27.6	29.6	– 7.2	10.2	32.6
Group C – D	35	35	6.7	28.6	43	–11.5	8.2	33.2

TABLE IV

Year XIV

Name of test	Percentages of children passing the tests					
	Vocabulary	Induction Test	President and King	Problems of Fact	Arithmetical Reasoning	Clock
Total group.....	18.4	49	20.6	45.5	40	21.8
Group A.....	37.3	48.5	28.4	57	40.3	22.4
Group B.....	7	50	16.3	38	41	21.4
Group C.....	37.3	40	27.4	60.8	35.3	11.8
Group D.....	6.6	54	14.5	41	46	22.3
Group A − B...	30.3	− 1.5	12.1	19	− 0.7	1
Group C − D	30.7	−14.0	12.9	19.8	−10.7	−10.5

TABLE V

Year XVI

Name of test	Percentages of children passing the tests					
	Vocabulary	Interpretation of Fables	Difference, Abstract Words	Problem of Enclosed Boxes	Repeat Six Digits Backwards	Code
Total group	2.4	12.7	5.5	30	26	11
Group A.....	6	21	12	34.3	28.4	20.7
Group B.....	0	7	1	26.5	24.5	5.2
Group C.....	6	21.6	11.8	33.3	25.5	19
Group D.....	0	9.2	1.3	30.3	28	6.6
Group A − B..	6	14	11	7.8	3.9	15.5
Group C − D ..	6	12.4	10.5	3	−2.5	12.4

21.1 is found as the coordinate of the column headed "Comprehension." This 21.1 is the percentage of children in Group A passing the "Comprehension" Test in excess of the percentage of children in Group B passing the same test.

For the purposes of further comparison the percentages of children passing the tests were averaged separately for each of the four years and for each of the five groups, A, B, C, D, and the total group. These averages were then summed and divided by four to give a total average

percentage of tests passed by each of the groups. To illustrate, take Total group, Year X. The percentages of children passing the tests are 100, 96.4, 87.8, 97, 80, and 78.8 respectively for the six tests that make up Year X. The average of these six percentages is 90. So, also, for years XII, XIV, and XVI, with average percentages of 61.5, 32.6, and 14.6. Summing these four averages and dividing by four gives 49.7, the total average percentage of tests passed by this group. Like averages were computed for the group of tests designated the Non-language group, and for the group of tests designated the Language group. These averages are given in Table VI. These same facts may be expressed in terms of months of mental age and are given in Table VII. The numbers of Table VII represent months of mental age and are averages for the several groups. They were determined by totalling the months "passed" on the tests of years X, XII, XIV, and XVI. This procedure may be readily understood by thinking of the tests used in the Non-language group as "starred" tests.

It will be seen from Tables VI and VII that for groups A and C there is fairly consistent performance on all three groupings of the tests, the performance on the Language group being in each case slightly higher than on the Non-language group. For groups B and D there is wide variation between the performance on the Language and Non-language groupings of the tests, and this variation holds proportionately for the total group of tests. This variation can best be expressed perhaps in terms of months of mental age. For Group B the average performance on the Non-language group of tests was 38.1 months, while the like performance on the Language group of tests was 24.4 months, or 13.7 months less. The variation for Group D was still wider, amounting to a difference of 17.3 months. When the performance of groups B and D on the Non-language group of tests is compared with their performance on the total group of tests, the difference is less marked. This is to be expected if there is a language handicap operative, as the total group of tests includes the Non-language group. An estimate of the amount of language handicap operative for Group B on the total group of tests for the years X, XII, XIV, and XVI may be arrived at by taking the difference between the number of months of mental age passed on the total group of tests and the number of months of mental age passed on the Non-language group of tests. This difference is 5.7 months. The like difference for Group D is 8.1 months. A slightly more refined estimate of the amount of language handicap is the difference in performance of groups A and B, and C and D, on

TABLE VI

	Total average percentage of tests passed		
	All tests of the four years	Non-language group	Language group
Total group	49.7	52.4	44.8
Group A	57.1	55.3	57.0
Group B	44.5	50.4	36.2
Group C	55.1	50.7	56.8
Group D	44.7	52.1	34.7
Group A − B	12.6	4.9	20.8
Group C − D	10.4	−1.4	22.1

TABLE VII

	Total months of mental age on years X, XII, XIV, and XVI		
	All tests of the four years	Non-language group	Language group
Total group	36.9	39.8	32.5
Group A	43.3	42.2	43.2
Group B	32.4	38.1	24.4
Group C	41.5	38.2	43.2
Group D	32.4	40.5	23.2
Group A − B	10.9	4.1	18.8
Group C − D	9.1	−2.3	20.0

the total group of tests minus their respective differences of performance on the Non-language group of tests. These differences are expressed in Table VII by the differences between the numbers 10.9 and 4.1, and 9.1 and −2.3, in the rows headed Group A − B and Group C − D, respectively. They take into account the superiority of performance of Groups A and C on the total group of tests over their

performance on the Non-language group of tests. The differences are, for Group B, 6.8 months and for Group D, 11.3 months. As some of the children taking the tests reached back into Year IX and others reached forward into Year XVIII, these estimates are slightly too low.

The Code Test of Year XVI offered considerably more difficulty to Group B than to Group A. It was included in the Non-language group because it was thought to be representative of performance requiring but little language ability, and because it would add to the varied nature of the Non-language group of tests. Two language factors enter into the taking of this test and perhaps account in part for the difference of performance of Groups A and B. One is the fact that the instructions to the child for taking this test are long and wordy, and the other is that the correct performance of the test involves the ability to spell the code words "come quickly." Each of these factors it is believed operated somewhat more against Group B than against Group A.

In giving the tests it was felt that the fables offered special difficulty to children from the foreign language-speaking homes. To test this impression, the individual scores made on the set of fables were averaged for each of the groups.

These averages follow. Total group 3.73, Group A, 4.87; Group B, 3.0; Group C, 4.92; and Group D, 3.1. The average number of words correctly passed on the vocabulary at the back of the booklet was also computed for each of the groups. These averages were: Total group, 42.5; Group A, 46.5; Group B, 39.3; Group C, 46.1; and Group D, 39.1. The fables offered approximately 40 per cent more difficulty for Group B, relative to the performance of Group A, than did the vocabulary.

The conclusion reached by the study is that there is a language handicap of from six to eight months of mental age on the test for children 13 years old coming from foreign-language-speaking homes. This conclusion is supported by what one would expect of children coming from homes where the parents are of the intellectual status that the parents of these children evidently are. The occupations of the fathers of these children mark them as of good average intelligence or above. Further, the fact that selection has already been operative to a considerable extent by the time children reach Grade VII of the public school is strong argument in support of the contention that the average intelligence of Grade VII school children is equal to the average intelligence of unselected children of like age. When children with a

language handicap are associated in school with children from English speaking homes, the forces of selection are likely to operate more rigorously against the children with the language handicap than would be the case were the children not so associated. It is the opinion of the writer that Group B of this study represents the more highly selected of the two groups.

STANDARDIZED CONTENT IN HIGH SCHOOL BIOLOGY

G. M. RUCH

State University of Iowa

AND

LEO H. COSSMANN

University of Oregon

The present report deals with a preliminary attempt to survey the content of high school courses in biology throughout the United States. Biology presents, perhaps, the least uniformity in subject matter today of any of the sciences commonly included in the secondary curriculum. In the main this is a decidedly wholesome situation although doubtless some effort toward the selection of a minimum, basal content would have the effect of raising the standards of instruction in biology in the smaller schools. The urgent need is for rigid experimentation as a foundation for a semi-standardized content for such courses. In the absence of such objective data, the second best means of arriving at the desired end is that of surveys of what the best schools are teaching in biology. To do this properly we might have recourse to several methods of investigation, *viz.* Analysis of textbooks, analysis of courses of study, questionnaire studies, or compilations based upon final examinations which have actually been given over a period of years in leading schools of secondary rank. The advantages among these possible methods favor, in the main, the last mentioned since the final examination represents a more or less thoughtful attempt on the part of the teacher to isolate those facts and principles which can with most reason be expected to be retained by the pupil for some appreciable length of time after the completion of the formal work of the classroom and laboratory. It is for this reason that the authors chose to attempt to survey good teaching practice in biology by statistical analysis of final examinations collected from all parts of the country. The limitations of the method are as apparent as are its advantages and no claim is made that finality of content determinations can or should be based upon this method.

Believing that the results of such a provisional investigation would possess enough merit to justify the effort expended, the authors[1]

[1] Credit for most of the actual work is due to Leo H. Cossmann.

began in 1918 actual work on the problem the completion of which has, for many reasons, been sadly delayed. The method of attack will be very briefly presented in summarized form.

1. A letter was sent to the Superintendent of Public Instruction in every State in the Union having a central officer asking that he list the names of those schools and teachers of his state who were doing the highest type of work in high school biology. Lists were received from 23 states with a total of 126 schools answering to our specifications.

2. Each of the 126 teachers was asked to send duplicate copies of all final examination questions used during the current year. Lists for previous years were also requested and in some cases were available for periods of five or more years prior to 1918. The total number of questions received ran slightly over 2000.

3. The 2000 questions were then analyzed for frequency and condensed and classified under convenient headings as 300 constantly recurring items. The group headings finally numbered 11 such phases of the total subject as: Definitions of terms, identification of structures, life histories, diagrams, functions, economic and ecological facts, etc.

4. The final list of 300 questions was submitted to about 100 of the leading teachers and authorities on the teaching of biology for ratings as follows:

(*a*) "Rate "1" all questions which you consider to be entirely satisfactory and representative of the kind of knowledge which should be presented to high school students of biology, and which high school students can reasonably be expected to answer.

(*b*) Rate "2" all questions which are partially satisfactory, but which you consider to be not highly desirable for any reason . . . , etc.

(*c*) Rate "3" all questions which are entirely unsatisfactory."

Provision was also made for the writing in of any other questions not already listed. Sixty-eight teachers and 9 authorities (authors of texts, professors of the teaching of science, etc.) rated all of the questions submitted. This was a remarkably high percentage of responses and whatever selection did result from the failures to answer can safely be assumed to be a selection upwards and hence could not have operated to decrease the reliability of the results.

Space will not permit the tabulation of all of the data so rated. Only the most significant portions of the total determinations will be presented and the lists will be further restricted to those items rated

"1" by the 77 judges although there is doubtless a great deal of value to be had from the analysis and tabulation of the "2" ratings.

It is in place at this point to refer to a second motive in carrying out this investigation. The authors hoped in advance (and this hope had been realized) to use the results in the production of at least two equal forms of a test for measuring accomplishment in high school biology. These tests are practically ready for publication and will be described in some detail in a later article.[1]

Tables I to II which follow present the ratings of the two sets of judges together with very greatly condensed statements of the items rated. Two numbers occur at the right of each item. The first is the actual number of "1" ratings given by the 68 teachers and the second is the actual number of "1" ratings from the 9 authorities. The notation "68:9" means, therefore, that the items was rated "1" by every judge in both groups, and similarly for ratings of less than perfect agreement.

TABLE I.—ITEMS DEALING WITH DEFINITIONS, LOCATIONS, IDENTIFICATIONS, USES, AND DIAGRAMS OF TERMS

1. Organ	68:9	26. Root hairs	67:8
2. Excretion	68:8	27. Protoplasm	67:9
3. Pistil	68:9	28. Immunity	67:9
4. Reproduction	68:9	29. Heredity	67:9
5. Oxidation	68:9	30. Ovule	67:9
6. Pollination	68:9	31. Crop rotation	67:9
7. Fertilization	68:9	32. Carbohydrate	67:9
8. Cross-pollination	68:9	33. Self-pollination	66:9
9. Parasite	68:9	34. Environment	66:9
10. Biology	68:9	35. Fermentation	66:7
11. Cell	68:9	36. Osmosis	66:8
12. Embryo	68:9	37. Larva	66:9
13. Protozoa	68:9	38. Tissue	65:8
14. Seed	68:9	39. Antitoxin	65:9
15. Respiration	68:9	40. Variation	65:8
16. Ovary	68:9	41. Pupa	65:8
17. Chlorophyll	68:9	42. Struggle for existence	65:8
18. Digestion	68:9	43. Thorax	64:9
19. Sterilization	68:9	44. Vertebrate	64:9
20. Pollen	68:9	45. Cilia	63:8
21. Cotyledon	68:7	46. Gametophytes	63:4
22. Grafting	68:7	47. Pasteurization	63:9
23. Evaporation	68:8	48. Transpiration	63:8
24. Sperm cell	68:9	49. Spore	63:9
25. Survival of fittest	67:9	50. Corolla	63:8

[1] The Ruch-Cossmann Biology Test, World Book Co., Yonkers, N. Y.

51. Enzyme	63:7	99. Bud scales	54:6
52. Fruit	63:9	100. Lower epidermis	54:6
53. Calyx	62:9	101. Lenticels	53:6
54. Assimilation	62:7	102. Saprophytes	53:6
55. Fungus	62:9	103. Coelenterates	53:5
56. Fibro-vascular-bundle	62:6	104. Fehling's solution	53:7
57. Irritability	61:6	105. Heliotropism	53:5
58. Metamorphosis	61:8	106. Gamete	52:3
59. Conjugation	61:7	107. Petiole	52:7
60. Ventricle	60:7	108. Rabies	50:8
61. Budding	60:6	109. Photosynthesis	50:9
62. Antennae	60:8	110. Erosion	50:5
63. Annuals	60:6	111. Ptomaine	50:8
64. Biennials	60:6	112. Mycelium	50:5
65. Perennials	60:6	113. Symbiosis	50:6
66. Palisade cells	60:5	114. Glottis	49:5
67. Epidermis	60:6	115. Hilum	49:6
68. Algae	60:5	116. Narcotic	49:7
69. Protective coloration	60:7	117. Zygospore	49:4
70. Element	60:8	118. Micropyle	49:6
71. Mimicry	60:7	119. Neuron	48:5
72. Hypocotyl	60:7	120. Gill rakers	46:6
73. Stimulant	60:7	121. Univalve	43:4
74. Cambium	60:7	122. Regeneration	43:5
75. Carnivorous animal	60:8	123. Vaso-motor system	43:4
76. Hybrids	59:7	124. Dehiscent	41:5
77. Humus	59:8	125. Anabolism	41:3
78. Combustion	58:8	126. Zoöspore	41:5
79. Warning coloration	58:7	127. Flagellum	41:4
80. Herbs	58:5	128. Alternation of generations	40:4
81. Monocotyledenous seed	58:6	129. Metabolism	40:5
82. Dicotyledenous seed	58:6	130. Contractile vacuole	40:6
83. Bud scars	58:6	131. Eugenics	40:6
84. Endosperm	58:5	132. Artificial propagation	40:7
85. Proboscis	57:6	133. Morphology	40:4
86. Chrysalis	57:8	134. Pseudopodia	40:4
87. Tropism	56:5	135. Incomplete flower	40:3
88. Poison	56:7	136. Chromosome	40:4
89. Lymph	56:8	137. Rhizoid	40:4
90. Rodent	56:8	138. Chordate	39:4
91. Cerebellum	55:7	139. Notochord	39:4
92. Bulb	54:7	140. Pelvic fins	38:5
93. Spleen	54:7	141. Tube feet	38:4
94. Shrubs	54:5	142. Clavicle	38:4
95. Symmetry	54:6	143. Incisors	38:8
96. Dominant	54:6	144. Convolutions	38:5
97. Mutation	54:6	145. Water vascular system	38:5
98. Diastase	54:8	146. Mesentery	37:5

147. Papillae	37:4	161. Operculum	26:4
148. Angiosperm	36:3	162. Pathenogenesis	26:3
149. Collar cell	35:4	163. Hermaphroditism	26:2
150. Cleistogamous flower	33:3	164. Trichinosis	24:3
151. Mitosis	33:1	165. Madreporic plates	23:3
152. Geotropism	30:5	166. Caeca	23:4
153. Spiracles	30:6	167. Chelipeds	21:3
154. Maxillipeds	30:5	168. Cuvier	19:2
155. Labial palps	30:4	169. Coppice woodland	19:2
156. Balancers	29:4	170. Ambulacral groove	19:2
157. Centrosome	28:4	171. Megalops	15:2
158. Pectoral fins	28:7	172. Acetanilide	11:6
159. Mediastinum	27:0	173. Glochidium	10:1
160. Blastula	27:5	174.[1] Lepidoptera	10:0

Etc.

TABLE II.—LIFE HISTORIES OF VARIOUS FORMS

1. Frog	65:9	10. Moss plant	52:3
2. Bee	65:8	11. Mold	52:4
3. Fly	62:9	12. Moth	49:6
4. Butterfly	62:8	13. Fern	49:3
5. Yeast	60:6	14. Parasitic worm	47:5
6. Grasshopper	57:8	15. Crayfish	47:4
7. Fish	57:7	16. Bracket fungus	39:1
8. Malaria organism	54:7	17. Starfish	36:1
9. Mosquito	52:9	18. Sponge	28:2

Etc.

TABLE III.—COMPARISONS OF FORMS AND STRUCTURES

1. Vertebrates and invertebrates	65:9
2. Distinctions between moths and butterflies	62:3
3. Differences between insects and spiders	62:4
4. Comparisons of reproduction in fishes and flowering plants	62:7
5. Differences between Reptilia and Amphibia	60:5
6. Distinctions between beetles, grasshoppers, bugs, and flies	60:3
7. Comparisons of arteries and veins	60:5
8. Internal structures of mono- and di-cotyledenous seeds	58:4
9. Numbers of eggs laid by birds and fishes; with reasons	57:7
10. Comparisons of food-getting and locomotion in Protozoa and insects	57:5
11. Likenesses of the grasshopper and butterfly	57:7
12. Respiratory mechanisms of birds, frog, and grasshopper	57:6
13. Comparisons of frogs and fishes in appearance, means of protection, respiration, reproduction, locomotion, importance	54:6

[1] The lower ranges of all the distributions have been omitted in Tables I to II for purposes of economy of space and because of the lesser utility of such information. The effect of this practice on the correlations computed will be considered at the end of the paper.

14. Comparisons of Protozoa and bacteria 54:6
15. Comparisons of saprophytes and parasites 50:5
16. Comparisons of structures of lowest animal and plant forms 49:3
17. Larval and embryonic development 49:3
18. Comparison of Paramoecium and frog for locomotion, food-getting, and digestion 47:6
19. Comparisons of Paramoecium and Amoeba 47:4
20. Comparisons of legs of grasshopper, bee, and fly 47:6
21. Properties of carbon, sulphur, and phosphorus 46:5
22. Differences between earthworm and coelenterates 45:1
23. Differences between living and non-living materials 45:8
24. Comparisons of plant and human nutrition 45:8
25. Comparisons of circulation of the frog and the fishes 45:5
26. Methods of reproduction in the black mold and Spirogyra 43:0
27. Comparisons of Protozoa and Metazoa 41:5
28. Comparisons of respiratory system of earthworm and man 41:0
29. Comparisons of breathing of birds with other animals 41:6
30. Comparisons and diagrams of digestive systems of man, frog, and fishes. 41:5
31. Comparisons of mouth parts of the bee and grasshopper 40:4
32. Nutrition of the molds, yeasts, and Spirogyra 39:3
33. Digestive systems of the hydra, honey-bee, and earthworm 36:2
34. Nutrition in Mucor with that of Spirogyra 34:1

Etc.

TABLE IV.—DIAGRAMS AND DRAWINGS WITH FULL LABELS

1. Typical cell 68:9
2. Complete flower 65:8
3. Essential organs of flowers 62:8
4. Germination stages of corn and bean 57:7
5. Typical insect 57:6
6. Heart and lungs 54:5
7. Woody stem 52:6
8. L. S. of bean and corn (seed and fruit) 52:6
9. Paramoecium 50:8
10. Amoeba 49:7
11. Diagrams of Mendel's laws 47:4
12. Hind leg of grasshopper and bee 41:5
13. L. S. of tooth 41:7
14. Habit of mushroom 39:3
15. Carbon cycle 36:5
16. Moss plant 36:2
17. Head of butterfly 34:3
18. Breathing apparatus of fish 36:4
19. Head of honey-bee 26:3
20. Nuclear changes in fertilized frog egg 20:0
21. Structure of skin 20:3
22. Maturation of sperm 18:0
23. Chick embryo at 36, 48, and 72 hours 15:0
24. Internal structure of crayfish 15:3
25. Different types of feathers 10:0
26. Digestive tracts of bird and clam 10:1

Etc.

TABLE V.—ADAPTATIONS OF FORMS

1. Adaptations of human body against disease germs 65:9
2. Definition of adaptations 65:7
3. Adaptations of flowers for attracting insects 63:9

4. Protective adaptations of fishes..... 62:7
5. Adaptations of flowers for prevention of self-pollination..... 62:8
6. Ecological adaptations of birds..... 60:7
7. Adaptations of fishes for locomotion..... 58:7
8. Adaptations of feet and beaks in birds; examples, and reasons..... 53:8
9. Mouth parts of bee and grasshopper for food-getting..... 52:7
10. Adaptations of proboscis of butterfly..... 52:7
11. Adaptations of frog for feeding, swallowing, and escape from enemies... 51:6
12. Adaptations of hind legs of bee or grasshopper..... 49:6
13. Adaptations of clam and bird for locomotion, protection, and food-getting..... 49:6
14. Adaptations of hydra for capture and intake of food, elimination of waste materials, and reproduction..... 41:3

Etc.

TABLE VI.—CLASSIFICATIONS

1. Name three orders of insects; with two examples of each..... 52:4
2. Outline of great plant groups with examples of each..... 49:1
3. Names and characteristics of nine of animal groups (phyla)..... 47:3
4. Classification of sponge by phylum name..... 47:1
5. Classification of following forms: Amoeba, toad, crab, slug, seagull, chicken, dragon fly, Culex mosquito, locust, moth, bee, mussel, scorpion, robin, salmon, man, frog, snail, clam, butterfly, house-fly, perch, centipede..... 47:5
6. Give the classification of the Arthropoda with characteristics of each... 47:1
7. The two great divisions of the seed-bearing plants, and four divisions of spore-bearing plants..... 44:3
8. Classes of invertebrates with examples..... 43:1
9. Classes of mammals with examples..... 38:1
10. Classes of fishes with examples..... 33:1
11. Divisions of the Gymnosperms (as detailed as possible)..... 20:0

Etc.

TABLE VII.—ECONOMIC IMPORTANCE; BENEFICIAL OR INJURIOUS

1. Bacteria..... 68:9
2. Forests..... 68:9
3. Mosquito..... 68:9
4. Birds..... 68:9
5. Yeasts..... 68:9
6. Fly..... 68:9
7. Potato beetle..... 65:9
8. Fish..... 64:9
9. San Jose Scale..... 62:9
10. Fungi..... 62:7
11. Bee..... 61:9
12. Ladybird beetle..... 60:9
13. Bumble bee..... 59:7
14. Molds..... 57:8
15. Grasshoppers..... 57:8
16. Earthworms..... 52:9
17. Moths..... 52:7
18. Butterflies..... 52:8
19. Mollusks..... 50:6
20. Mosses..... 48:4
21. Lichens..... 46:4
22. Sponge..... 44:6
23. Frogs..... 44:6
24. Ferns..... 39:3
25. Starfish..... 35:4
26. Squid..... 33:2

TABLE VIII.—HISTORY OF BIOLOGY (EVOLUTION)

1. With what facts are the following names associated?

Darwin	68:8	Lamarck	52:4
Mendel	65:8	Jenner	51:6
Pasteur	65:8	Harvey	49:6
Burbank	62:7	Lister	47:5
De Vries	60:1	Schleiden and Schwann	44:3
Koch	54:5	Metchnikoff	41:4

Etc.

2. What is the theory of evolution: Who is the originator? 62:9
3. What are the bearers of heredity? 47:6
4. What part did geology play in the development of the theory of evolution? 39:6
5. State four ways of accounting for the evolution of plants and animals .. 39:4
6. Trace the development of the nervous systems in the animals you have studied 37:4
7. What is the Law of Recapitulation? 35:3
8. Why do scientists believe that birds have descended from reptiles? 35:4
9. What transition family remains between the invertebrates and vertebrates? 31:2
10. What did the Echinoderms contribute to the development of animals? . 27:1
11. Trace the evolution of the heart from the fishes to the mammals 22:3

TABLE IX.—GENERAL DESCRIPTIONS, DISCUSSIONS, AND EXPLANATIONS; LARGELY OF ECONOMIC AND HYGIENIC IMPORTANCE

1. State how the following measures aid in controlling the spread of epidemics: (*a*) quarantine, (*b*) vaccination, (*c*) boiling drinking water, and (*d*) school nurses 68:9
2. Discuss importance of forests as to: (*a*) climate, (*b*) soil, (*c*) erosion 68:8
3. Discuss the following:

Disposal of garbage	68:9	Dangers of patent medicine	63:8
Reasons for cooking food	68:9	Preservatives	63:8
Sources of contamination of food and water	68:9	Coughing and sneezing	62:9
How animals carry disease	68:7	Public parks	61:8
Diseases caused by bacteria	66:9	Effects of alcohol on physical endurance	60:8
Disinfectants	65:9	Treatment of sprains	57:7
Bacteria and agriculture	65:8	Bacteria and dairying	57:7
Oiling of streets	65:8	Prevention of spinal curvature	54:8
Tobacco and scholarship	65:8	Nitrification and denitrification	51:4
Noise and mouth breathing	65:9	Soil fertility and improvement	49:6
Use of the common drinking cup	65:9		

Etc.

4. Explain the following:

How the leaf does its work	65:7	Inheritance of feeble-mindedness	54:6
Methods of seed dispersal	63:8	Rules of hygiene of the circulation, respiration and digestion	53:6
Necessity for seed dispersal	63:8	Composition of human blood	52:9
Meaning of "Environment"	63:8	Crop rotation	52:8
Principles of ventilation	60:7	Nitrogen cycle	51:7
Inter-dependence of plants and animals	60:7	Why the Paramoecium is a "higher" animal than the Amoeba	49:5
Development of the seed from the ovule	58:7	Most economical methods of increasing soil fertility and productivity	49:6
Relation of Protozoa to disease	57:8		
Food manufacture in plants	57:7		
Formation of habits	54:6		

Etc.

5. Describe the following:

Spirogyra	49:3	External and internal structure of the bird	36:4
Community life of the ants	49:4	Stages in manufacture of vinegar from apples	35:5
Structure of human brain	44:3	Sponge in detail	20:3
One of the crustaceans	41:4		
One of the fishes	41:5		
An ant colony	39:4		

Etc.

TABLE X.—FUNCTIONS OF STRUCTURES AND PROCESSES

1. Leaves	68:9	11. Chloroplasts	62:7
2. Essential organs of plants	68:9	12. Organs of circulation in man	61:7
3. Roots	68:6	13. Plasma, lymph, thoracic duct	60:8
4. Fruits	68:7	14. Proboscis of butterfly	60:7
5. Root hairs	68:8	15. Digestive glands of man	57:8
6. Stomata	68:6	16. Parts of bean seed	55:6
7. Circulation of the blood	68:6	17. Functions common to all animals	51:5
8. Blood	68:8	18. Lenticels	47:5
9. White and red blood corpuscles	65:8	19. Hind feet of honey bee	43:5
10. Functions of protoplasm	62:8	20. Parts of brain	42:7

Etc.

TABLE XI.—UNCLASSIFIED MATERIALS

1. What should be the farmer's attitude toward the following: Snakes, owls, hawks, cats, crows, wasps, poisoning of gophers, blackbirds, ichneumon flies, and kingbirds?	65:8
2. Name two foods rich in protein, grape sugar, and starch	65:9
3. Name important excretions of plants and animals	65:5
4. Name the most effective ways of killing bacteria	64:8
5. Name three economical foods rich in fuel value and three foods rich in building materials	63:9

6. Tests for starch, protein, grape sugar, and fats 61:8
7. Name the most common chemical elements in foods 61:9
8. Name the elements used by green plants in food manufacture 61:8
9. Name the chief nutrients in a bowl of bread and milk 60:9
10. State the characteristics of each of the following and state uses to living forms: Carbon, oxygen, nitrogen 60:8
11. What are five essential conditions for plant growth? 57:6
12. Fill out the following table with reference to disease: 57:8

Disease	Organism Concerned	Exact Mode of Infection	Best Preventive Measures
Diphtheria			
Smallpox			
Tuberculosis			
Typhoid fever			

13. Name and give the characteristics of four chemical elements 55:8
14. Give six methods of propagation in plants 54:6
15. Name five nutrients and their uses to the human body 54:6
16. How is the heart affected by violent exercise? 53:6
17. Name three ways in which the heart is protected by the structure of the human body 53:6
18. What two kinds of muscles are found in the human body? 54:6
19. What is the chief use of carbon in the human body? 54:6
20. State number, position, and kinds of teeth in man 51:8
21. Does alcohol tend to increase feeble-mindedness? 51:8
22. What are the proper proportions for a standard daily diet? 50:8
23. Name three processes of fermentation, stating the organism responsible, and the substances acted upon 48:3
24. Name five tissues of the human body and characterize each 44:4
25. Give in tabular form: (*a*) the digestive juices, (*b*) the enzymes found in each, and (*c*) the action of each enzyme 43:5
26. State whether the following are physical or chemical changes:
    Action of diastase on starch
    Dissolving of sugar
    Osmosis of sugar
    Oxidation of sugar
    Dessication of spores 39:5
27. What are the specific effects of muscular exercise? 34:4
28. Outline the divisions of the human skeleton and name the bones in each division 28:3
29. Name the principal tissues formed from the middle layer of the gastrula 26:1

Etc.

Attention should be specifically drawn to the fact that in each table which has been presented the lower ranges have been eliminated. No systematic rule was observed in this arbitrary elimination of the lowest ratings. In general, the printed tables cover from 50 per cent to 75 per cent of the total range, the average being not far from 70 per cent.

Certain of the groups of items have been much more drastically curtailed than others upon a basis of rough inspection of relative merit. The only justification for this procedure which is offered reduces to questions of limitation of available space.

Although inspection of the tabulations raises many very significant questions, only a very few can be discussed here. In Table VI, for example, there is a very evident tendency for the group of professional experts to minimize the importance of the taxonomic relations. But one of the questions dealing with classification is endorsed by more than 50 per cent of the professional judges in comparison with the fact that 9 of the 11 of these items are selected by more than 50 per cent of the teacher group. This finding is quite in harmony with the literature on the teaching of biology during recent years and is probably significant evidence that systematic problems are gradually being eliminated from our courses in biology. Similar tendencies against the teaching of biological technicalities among the professional students of the problem can be found in other tables, *e.g.*, Table IV. In general the professional group can be shown to favor economic and practical problems in high school science.

A final question which must be answered statistically is the matter of the general reliability of the judgments obtained. The reliability coefficients appearing in column (*a*) of Table XII which follows were obtained by correlating the numbers of "1" judgments given by the professional judges with those of the teachers; the numbers being 68 and 9, respectively. Column (*b*) gives the probable error of the coefficients (Pearson Product-moment formula). Column (*c*) gives the reliabilities for the sums of judgments in both groups as estimated by Brown's formula:

$$r(68 + 9)(68' + 9') = \frac{nr}{1 + (n - 1)r}$$

The interpretation of these reliability coefficients is very difficult since they are computed from the actual judgments recorded in the tables given above rather than from the complete data which included the lowest 25 per cent to 50 per cent of the range as well. This serious curtailment of the range and hence of the standard deviations will of necessity cause the reported coefficients to fall markedly short of the true values based upon the full range. As would be expected the reliability coefficients are lowest in those tables which are most abbreviated in the lower ranges. We can conclude without much danger of

error that all of the reliabilities are moderately or very high, particularly for the data in Tables I, II, IV, VII, and IX. Even in the cases of Tables III and X which appear to be very low, it will be noticed that hardly 50 per cent of the total range was reported and used in the computations. Upon the assumption that the true sigma for the whole data is nearly twice the size of the sigma for the reported data in any of the distributions, it can be prophecied from Kelley's formula

$$\frac{\sigma_t}{\sigma_T} = \frac{\sqrt{r(1-R)}}{\sqrt{R(1-r)}},$$

that, upon the assumption that $\frac{\sigma_t}{\sigma_T} = \frac{1}{2}$, the prophecied R's will, with one exception, all exceed 0.80. The validity of this assumption concerning the true range can be roughly tested by inspection of the tabulated results. The matter of reliability can properly be dismissed with the statement that the obtained values are about as high as are ordinarily secured by the method of estimates even without the corrections attempted, and that if the described corrections are applied as hypothecated, the reliabilities are unusually high in most cases.

TABLE XII.—RELIABILITY OF THE JUDGMENTS

Table number	(*a*) Obtained r (Pearson)	(*b*) PE	(*c*) r for combined judgments estimated by Brown's formula
1	.80	.02	.82
2	.75	.08	.77
3	.35	.10	.38
4	.91	.02	.92
5	.72	.09	.74
6	.58	.13	.61
7	.79	.05	.81
8	.74	.07	.76
9	.83	.03	.85
10	.54	.11	.57
11	.58	.08	.61

MENTAL AGE AND SCHOOL ATTAINMENT OF 1007 RETARDED CHILDREN IN MASSACHUSETTS

PREPARED UNDER THE DIRECTION OF AMY HEWES
BY MILDRED HOLT, SOPHIE MERANSKI, JULIA SNELL

Mt. Holyoke College

The educational waste caused by having children of retarded mental development in the regularly graded classes has long been apparent. Both the slow children and those who learn rapidly suffer unless the course of study is adapted to their capacity. In 1919 the Massachusetts Legislature enacted a statute which requires that "the school committee of each city and town in which there are 10 or more children three years or more retarded shall establish special classes to give such children instruction adapted to their mental attainments."[1]

To carry out the provisions of the act, traveling clinics have been established for the examination of every retarded pupil in the State. These examinations are comprehensive and include a number of different fields of inquiry such as physical condition, family history, economic and social history, moral reactions, and practical knowledge, in addition to school attainment and mental age, which are the two findings especially dealt with here. The Stanford Revision of the Binet-Simon Tests was used to determine mental age, and especially prepared school tests to establish grade standing in the different school subjects. Experience in administering the Massachusetts law has emphasized the fact that the diagnostic significance of the comprehensive inquiry is very great, especially when used as a basis for advice and recommendations for individual children. A somewhat separate problem is that of ascertaining the grade of school work which the child is able to perform. In general the retarded children were "found in school grades two or three or more years above the grades whose work they were able to do."[2] An important question for the future school program, then, is: What correlation exists between the mental age level ascertained and the degree of educational proficiency as shown by the grade reached in the separate examinations in the different educational subjects. This is an inquiry which may have implications as well regarding the general appropriateness of different

[1] Massachusetts General Acts, Chapter 277, Approved July 1, 1919, Sec. 2.

[2] Fernald, W. E.: The Inauguration of a State-wide Public Mental Clinic in Massachusetts. *Mental Hygiene*, Vol. 6, July, 1922, p. 3.

subjects in the curriculum and possibly also the effectiveness of the pedagogical methods employed in each.

The Material Used.—The records on which the study is based were made from examinations of 1007 retarded children in 1922, and include the determination of mental age, intelligence quotient, and the proficiency attained in each of the following elementary subjects: Reading, arithmetic, spelling, writing, geography, and language. They were taken from records made in cities in eastern Massachusetts and were drawn off from each city in the order in which the children happened to be examined. The records were plotted on charts to make easy comparison with the corresponding school performance of the normal child.

In these charts the mental ages and grades which indicate the usual performance of those ages are arranged in adjacent parallel columns. Thus the examiners assume that a child who has a mental age of six can ordinarily do Grade I work, one with a mental age of seven, Grade II work, so that in most cases the school performance in the separate subjects is indicated to the right in a more or less straight line roughly corresponding to the appropriate grade. The records were tabulated to show the distribution of children according to mental age and school attainment in each subject, and the Pearsonian Coefficient of Correlation was computed.

TABLE I.—DISTRIBUTION OF 1007 RETARDED CHILDREN BY MENTAL AGE

Mental age in years	Total	Girls	Boys
Total	1007	357	650
4 and less than 5	2	2	
5 and less than 6	5	1	4
6 and less than 7	158	65	93
7 and less than 8	274	110	164
8 and less than 9	331	106	225
9 and less than 10	166	55	111
10 and less than 11	53	13	40
11 and less than 12	16	4	12
12 and less than 13	1	...	1
13 and less than 14			
14 and less than 15	1	1	

The Group Studied.—The group was composed of 357 girls and 650 boys. The chronological age varied from seven to seventeen years inclusive, and about half of the children were under thirteen years of age. The mental ages varied from four to fourteen years, and less than half (44.0 per cent) of the retarded children had a mental age of less than eight years (see Table I).

More than two-thirds (67.3 per cent) had an intelligence quotient of less than 70 and would ordinarily be classified "feeble-minded." The remaining third (32.7 per cent) included the difficult borderline cases and four pupils with intelligence quotients of 100 and over whose retardation was clearly due to cause other than mental incapacity (see Table II).

TABLE II.—DISTRIBUTION OF 1007 RETARDED CHILDREN BY INTELLIGENCE QUOTIENT

Intelligence quotient	Number of children
Total	1007
30 and less than 40	4
40 and less than 50	43
50 and less than 60	221
60 and less than 70	410
70 and less than 80	250
80 and less than 90	59
90 and less than 100	16
100 and over	4

The Significance of the Correlation.—The coefficients of correlation given below show the extent to which school performance in the various subjects corresponds to the mental age of the boys and girls tested. The correlation is high for all subjects, but highest of all for language, arithmetic, and writing, and lowest for geography.

SUBJECT	CORRELATION: MENTAL AGE AND ATTAINMENT IN SCHOOL TESTS
Language	.82
Arithmetic	.78
Writing	.72
Reading	.68
Spelling	.66
Geography	.59

The explanation of the differences in the degree of correspondence is not hard to find. The subjects in which the correlation is highest are those subjects *which are employed or paralleled in the mental tests themselves.* Language, for which the coefficient is highest, is, of course, the mechanism by which the child whose mental age is being determined expresses his capacity. In giving the definitions or solutions of problems required in the mental tests his rating rests in large part on his ability to express himself. Arithmetic, the second in rank, is itself used in the mental tests for older children, and also requires a close parallel to the reasoning processes involved in the answers to some of the questions. Writing, for which the correlation is also high, corresponds to still a third aspect of the mental tests; it requires very nearly the same motor coördinations as those involved in the copying problems used in the tests.

The lower correlation for geography is equally inevitable when the nature of the mental tests is recalled. These tests involve almost nothing of the pure memory work which geography is likely to be. That is, in this case the subject is *not* utilized in the very tests by which mental age was originally determined. Reading and spelling, which stand nearer the middle of the scale of correspondence, are less easy to classify. They are partially represented in the tests, but their use is not so frequently paralleled.

The closeness of correspondence between school performance and the mental age of defective children is, therefore, fore-ordained, in those school subjects which are themselves represented in the mental tests used. Miss Mary Moran, head-teacher in the Massachusetts School for the Feeble-minded writes: "In the majority of cases we can almost predict from the school tests what the mental age will be."[1] The danger of using this type of evidence as the basis of conclusions on the suitability or efficacy of any one type of mental tests is therefore apparent. The findings of the study on this point raise the same questions as are to be found in the criticism of psychologists who have discussed the determination of native intelligence by the tests.

The performances of retarded children show a considerable degree of variation among the school subjects. The extremes are spelling, in which a majority (60.7 per cent) failed to make the grade corresponding to their mental ages; and writing, in which more than one-half

(53.9 per cent) were above the grade to which their mental ages entitled them (see Table III).

TABLE III.—PERCENTAGE DISTRIBUTION OF RETARDED CHILDREN IN RELATION TO GRADE CORRESPONDING TO MENTAL AGE

Subject	Children in relation to grade corresponding to mental age		
	Per cent below	Per cent within	Per cent above
Spelling	60.7	25.8	13.5
Reading	56.0	29.8	14.2
Arithmetic	49.1	37.4	13.5
Geography	41.5	41.5	17.0
Language	33.2	36.8	30.0
Writing	15.1	31.0	53.9

The analysis of the records for the 1007 retarded children leaves the question open as to whether the high correlation of mental age can have any particular significance as long as the two sets of tests have common ground. Until the psychologists devise a test of intelligence which does not depend mainly on school performance their assistance for school grading is of little value.

The explanation of the differences in the degree of correspondence is not hard to find. The subjects in which the correlation is highest are those subjects *which are employed or paralleled in the mental tests themselves.* Language, for which the coefficient is highest, is, of course, the mechanism by which the child whose mental age is being determined expresses his capacity. In giving the definitions or solutions of problems required in the mental tests his rating rests in large part on his ability to express himself. Arithmetic, the second in rank, is itself used in the mental tests for older children, and also requires a close parallel to the reasoning processes involved in the answers to some of the questions. Writing, for which the correlation is also high, corresponds to still a third aspect of the mental tests; it requires very nearly the same motor coördinations as those involved in the copying problems used in the tests.

The lower correlation for geography is equally inevitable when the nature of the mental tests is recalled. These tests involve almost nothing of the pure memory work which geography is likely to be. That is, in this case the subject is *not* utilized in the very tests by which mental age was originally determined. Reading and spelling, which stand nearer the middle of the scale of correspondence, are less easy to classify. They are partially represented in the tests, but their use is not so frequently paralleled.

The closeness of correspondence between school performance and the mental age of defective children is, therefore, fore-ordained, in those school subjects which are themselves represented in the mental tests used. Miss Mary Moran, head-teacher in the Massachusetts School for the Feeble-minded writes: "In the majority of cases we can almost predict from the school tests what the mental age will be."[1] The danger of using this type of evidence as the basis of conclusions on the suitability or efficacy of any one type of mental tests is therefore apparent. The findings of the study on this point raise the same questions as are to be found in the criticism of psychologists who have discussed the determination of native intelligence by the tests.

The performances of retarded children show a considerable degree of variation among the school subjects. The extremes are spelling, in which a majority (60.7 per cent) failed to make the grade corresponding to their mental ages; and writing, in which more than one-half

---

[1] Moran, Mary F.: "School Tests for School Grading and Measures of Intelligence." Pamphlet, n. d.

(53.9 per cent) were above the grade to which their mental ages entitled them (see Table III).

TABLE III.—PERCENTAGE DISTRIBUTION OF RETARDED CHILDREN IN RELATION TO GRADE CORRESPONDING TO MENTAL AGE

Subject	Children in relation to grade corresponding to mental age		
	Per cent below	Per cent within	Per cent above
Spelling	60.7	25.8	13.5
Reading	56.0	29.8	14.2
Arithmetic	49.1	37.4	13.5
Geography	41.5	41.5	17.0
Language	33.2	36.8	30.0
Writing	15.1	31.0	53.9

The analysis of the records for the 1007 retarded children leaves the question open as to whether the high correlation of mental age can have any particular significance as long as the two sets of tests have common ground. Until the psychologists devise a test of intelligence which does not depend mainly on school performance their assistance for school grading is of little value.

THE SIGNIFICANCE OF WEIGHTED SCORES

PAUL V. WEST

University of Chattanooga

After a great deal of emphasis on the necessity of weighting of specific items or problems in examinations and tests, doubt is beginning to be manifested regarding the value of such emphasis. Douglass and Spencer[1] present a report of statistical treatment of results of certain standard tests, *i.e.*, Douglass Algebra, Henmon Latin, Gregory Language, and Monroe Silent Reading, and although no definite conclusion is stated, all the data presented are consistent in showing weighting to be of negligible value, with various groups ranging in size from 11 to 32 pupils in the Junior High School. Charter's contention that the correlation of weighted and unweighted scores is over .90 is found to be sustained.

The problem here noted arose in connection with the construction of a test of ability to reproduce the thought in passages read aloud by the examiner. Two tests of a similar nature were used, each consisting of five parts. In Test A the points of possible credit were: Part I, 11; Part II, 14; Part III, 20; Part IV, 11; Part V, 18; or a total of 74. In Test B the credit points were: Part I, 11; Part II, 19; Part III, 24; Part IV, 25; Part V, 21; or a total of 100. The two tests were given on different days to a group of 45 High School students of mixed grades from 8 to 12.

The question then arose as to whether the score on each test should be given as the aggregate of the simple points gotten on the test or whether each point, or item, should be weighted as to difficulty. Both methods were followed and results compared.

The process of weighting used was simply (1) Tabulation of results so as to show which items were passed satisfactorily by each member of the class. (2) Computation of the percentage of failure of the subjects on each item. (3) Translation of the values thus found into terms of "percentile" value by the use of the tabular form, p. 392 of Rugg "Statistical Methods Applied to Education." (4) Addition of these values, and computation of the percentage ratio that each item value held to the total, so that the sum of all these values would be equal to 100.

---

[1] Is it Necessary to Weight Exercises in Standard Tests: *Journal of Educational Psychology*, Vol. XIV, No. 2, Feb., 1923, pp. 109–112.

Each pupil's score was now found as follows: Pupil (a) passed successfully points 1, 2, 5, 8, 9, 11 in Part I of Test A giving an unweighted score of 6. But the weights found on these items were respectively .6, .6, .9, .8, .9, 1.4 giving a weighted score of 5.2. For the sake of convenience in interpretation all scores for each part as well as for the total were reduced to percentage values of the total possible score. For example: Subject (a) made a raw score of 6 points on Part I out of a possible 11 points giving him a percent score of 55. When the points were weighted he made a score of 5.2 out of a possible 10.3 or a percent score of 50.5.

Coefficients of correlation were computed by the Product-moment Method for the two sets of scores, raw and weighted, the scores being handled in discrete series. The results were unexpectedly high as shown in Table I. None of these coefficients shows a low enough relation to justify the conclusion that the weighting is greatly significant in changing the score or the ranking of any one of the individuals. The chances are greatly in favor of the simple point score being as correct as the weighted score.

TABLE I.—COEFFICIENTS OF CORRELATION BETWEEN RAW AND WEIGHTED SCORES FOR THE FIVE PARTS OF EACH OF TWO TESTS OF COMPREHENSION OF PASSAGES READ ALOUD TO STUDENTS

	PART I	PART II	PART III	PART IV	PART V	TOTAL
Test A.....	.991	.996	.987	.997	.993	.995
Test B.....	.995	.995	.992	.996	.998	.997

When one adds to this fact the difficulty of securing accuracy at every involved stage of the computations and the statistical manipulations requisite to the weighting procedure the reasons for weighting the items in this particular test are not easily found.

It is, however, possible that weighting may not be needed in one type of test while greatly needed in another type. The weighting process assumes a natural difficulty of certain items for the members of the group tested, so that only the very able will be able to pass the most difficult at all satisfactorily. But in the tests cited above much depended evidently on incidents of interest, attention and experience of the subject, so that the poor performer might pass in items which would give failure to the better worker and fail in those items which would prove very easy for the group as a whole.

Furthermore the number of items which are taken as the base in computing the percentage of failure will probably influence the weighting factor greatly. For example, in the above analysis, if the percent-

age of failure had been computed for each of the parts of the tests independently of any other part instead of grouping all of the points first, it is evident that the weights would be correspondingly more greatly differentiated.

Form 8 of the Army Alpha Test was given to the same group of pupils and the results for each of the eight test groups were analyzed and studied by the same process as before noted. The items of Test I (Directions) were scored by points as is customary. The 12 items in this test were then weighted, omissions and failures both being counted in the percentage of failure. In the Army Alpha Test each test is given with a definite time limit in contrast with the test previously noted, in which all subjects had all the time needed to give an answer to each question after it was orally presented following the reading.

Since the Alpha Test is an "intelligence" test one might anticipate that the element of inherent difficulty in certain elements of the tests would be more apparent and count for more in the weighting process. Especially might this be true since the items are scaled and the more difficult would tend to be omitted as well as wrongly answered to a greater degree.

Tests IV and V are those which require alternative responses and are scored by the Right Minus Wrong Method, consequently do not yield to weighting readily, at least by the above process. The question arises as to how the weightings should be applied so as to make allowance for guessing. The weightings were found as with other tests, but when individual scores were determined it was found that the errors often penalized unduly, and the results were discarded.

The remaining tests 1, 2, 3, 6, 7 and 8 were analyzed to note the effect of weighting. The weighted scores were correlated with the unweighted scores for each test with the results given in Table II. It is noteworthy that, although these coefficients are indeed high they are all lower than any of the coefficients in the case of the first tests worked with. Tests 3 and 8, are markedly lower than the others. This would seem to indicate either that the weighting of a smaller number of test items in a group gives greater significance to the weights or that certain tests make a greater demand for weighting, possibly due to the fact that they are more adequate measures of specific or general ability than others.

TABLE II.—COEFFICIENTS OF CORRELATION BETWEEN RAW AND WEIGHTED SCORES FOR CERTAIN TESTS OF THE ARMY ALPHA TEST

TEST 1	TEST 2	TEST 3	TEST 6	TEST 7	TEST 8
.975	.956	.932	.966	.984	.9402

With correlations of raw and weighted scores as high as was the case in Table I it is evident that the intercorrelations of the tests would not be greatly varied when the scores were weighted as compared with those from unweighted scores. But the Tests in the Army Alpha were correlated with each other, first by the use of raw scores, then by the use of weighted scores. In Table III the latter coefficients are shown in bold faced type. The marked agreement of many of the pairs of coefficients is striking as in the correlations of tests 1–6, 1–7, 1–8, 3–6, 3–8, 6–7. Some of the coefficients are lower, notably 1–2, 2–3, 2–7, 3–7, and 6–8, but just what this signifies is difficult to say, except that the factors affected by weighting are not such as to give the extreme

TABLE III.—COEFFICIENTS OF CORRELATION BETWEEN VARIOUS TESTS IN THE ARMY ALPHA TEST WHEN USING RAW SCORES AND WHEN USING WEIGHTED SCORES

Tests	1	2	3	6	7
2. raw	.601				
weight	**.504**				
3. raw	.333	.518			
weight	**.422**	**.469**			
6. raw	.507	.58	.50		
weight	**.481**	**.721**	**.499**		
7. raw	.587	.731	.673	.639	
weight	**.617**	**.673**	**.573**	**.643**	
8. raw	.426	.40	.542	.291	.485
weight	**.442**	**.453**	**.55**	**.23**	**.559**

cases any marked advantage, or agreement. The remaining tests alone show an increase of coefficient through weighting, and this is not a notably large change in any case. In the Army Alpha Test then we find greater evidence of the validity of the weighting process in point of its influence on the scores, and a clear example of the differentiation of tests in respect of demand for weighting, but it must be remembered that these data cannot determine whether the weighted or unweighted scores are the more reliable. In general the findings

here agree with the statement of Yerkes,[1] et al.: "Our original use of the sum of unweighted scores in all tests as the definition of the combined scale for the measurement of intelligence is not in error for all practical purposes in so far as the internal evidence of the system is relevant."

A third test was investigated, being applied to the same group of pupils, 44 cases. It was planned to use here a test which had had the items carefully weighted and scaled so that comparisons might be made between the simple unweighted scores, the weighted scores from the scaled values and the weighted scores as empirically determined.

The test used was the collection of 200 analogies as standardized by R. Pintner.[2] These analogies were carefully divided into five tests of 40 analogies each so that the accumulated weighted values of each test would be the same and each analogy in each test would be set over against others of approximately the same scaled value in the other tests.

The five tests were then applied to the pupils, without a definite time limit being used, and the results were scored by three methods; by the simple point score, by use of the weighted values assigned by Pintner, and by weightings as derived by the same method as in the preceding tests, each test group being weighted independently of the others.

Intercorrelations of the five tests were worked out for each of the three sets of scores and the results are presented in Table IV. It is noteworthy that none of the intercorrelations is greatly altered by the

TABLE IV.—INTERCORRELATIONS OF SCORES IN FIVE ANALOGY TESTS WITH THE USE OF THREE DIFFERENT METHODS OF SCORING

TEST PAIRS	UNWEIGHTED POINTS	SCALE OF PINTNER	EMPIRICAL WEIGHT
1–2	.897	.857	.829
1–3	.863	.846	.861
1–4	.733	.665	.621
1–5	.717	.764	.689
2–3	.755	.773	.746
2–4	.616	.516	.502
2–5	.627	.700	.574
3–4	.759	.687	.672
3–5	.753	.814	.731
4–5	.567	.533	.514

[1] Psychological Examining in the United States Army. *Memoirs National Academy*, Vol. XV, No. 3, p. 651.

[2] *Journal of Applied Psychology*, Vol. IV, Nos. 2 and 3, June-Sept., 1920, pp. 263–273.

method of scoring. In every case conclusions would be the same regardless of which set of scores was used. However, the correlations with the raw scores are found to be generally higher than those in which the weights are derived, showing that the element of difficulty involved in any analogy is not so common as to be detected by weighting.

A typical comparison is here given in Table V of Rankings of the Subjects in one of these tests by the three methods of grading. There is a marked agreement in the three sets of ranks, in fact the variations are largely due to the comparatively large groupings of frequencies in the first set as compared with the other two. The extreme cases, especially, are in very close agreement.

TABLE V.—SHOWING COMPARISON OF RANKINGS OBTAINED BY THREE DIFFERENT METHODS OF SCORING A TEST

(*A*) Unweighted; (*B*) Pintner Scale; (*D*) Derived weights

Subject number	Rankings			Subject number	Rankings		
	A	*B*	*C*		*A*	*B*	*C*
1	22½	18	22	23	30½	24½	22
2	30½	31	30½	24	4½	5	4½
3	2	2½	2½	25	22½	27	26
4	43	43	43	26	9	6½	7
5	9	9½	7	27	22½	16	14
6	14½	13	13	28	40	40½	39½
7	14½	17	17	29	2	1	1
8	14½	14½	17	30	42	42	42
9	22½	27	22	31	37	38	37½
10	9	9½	10½	32	2	2½	2½
11	22½	24½	28½	33	30½	32	32
12	41	40½	41	34	37	39	39½
13	9	9½	10½	35	22½	29½	22
14	22½	22	26	36	9	9½	10½
15	14½	14½	17	37	9	6½	7
16	9	12	10½	38	22½	19	17
17	22½	20	28½	39	34	33	34½
18	22½	22	22	40	37	35½	36
19	44	44	44	41	4½	4	4½
20	34	34	33	42	30½	29½	30½
21	22½	27	26	43	34	35½	34½
22	39	37	37½	44	22½	22	17

It is not to be doubted that weighting has a definite part to play in test construction and use, but it is quite as certain that this part must be more clearly defined than at present. It is possible to spend a great deal of time and effort in the computation and use of weighted scores with no adequate or comparable return in the improved differentiation of pupils in an average group. It is possible that weighting will serve its best purpose in assisting in a scaled arrangement of material, even though the items be not scored by different values, but this and other possibilities must be determined through a more rigid analysis of data and the thorough testing of fundamental assumptions involved in the process.

INFORMATION AS A MEASURE OF INTELLIGENCE AND MATURITY

WILLIAM R. WILSON

University of Washington

The difference between tests that set out to measure intelligence more or less directly by measuring adjustments to novel situations, and those tests that are wholly or in part tests of maturity or experience has been recognized for 20 years. It has been pointed out many times that tests designed to discover differences in intelligence are valid measures only when the individuals composing the group examined are homogeneous as to past training and familiarity with the materials that compose the test. Such tests as color naming, giving dates, definitions, vocabulary, spelling, giving the value of stamps and coins, and general information, are regarded with some suspicion as likely to be unduly influenced by mere length of residence in the world. When these suspected tests are given to children of different ages there is found a fairly uniform increase in ability to handle them adequately as age increases. What part of this increase in ability is due to growth of intelligence with maturation and what part is due simply to further experience is not known.

When subjects of the same age and experience are to be compared with respect to general mental ability, the inclusion of the information type of test is scarcely open to criticism, as mental maturity and information show characteristic concomitant variation from year to year. However, little of our group testing in schools is of this kind. In the larger part of mental testing in school the individuals compared are usually homogeneous as to school grade rather than as to age or experience, and are often not even of the same school grade. For example, a test of the general type of the Army Alpha is given to all the pupils of a four-year high school, and the resulting scores are used as a basis for the comparison of the pupils as to general ability or whatever function it is assumed that the test measures. These pupils range in age from 12 to 20 years, from childhood nearly to complete mental and physical maturity. The range of educational achievement is four years. In spite of these facts, on the average one-fourth of the questions attempted in the Alpha test are in the general information section (Test 8) where ability to score may be largely contingent upon maturity rather than upon native ability. In comparing the Alpha test grade of a child of 12 with a man of 20, what allowance should be made for this fact? Or, more generally, if the Alpha test is given in a high

school what correction, if any should be made for the predominant weight given to the information test?

In the study of the problems thus suggested the test papers of 168 pupils of Grades IX to XII in the Grand Rapids (Minnesota) High School have been used. The tests (Alpha, Forms 6 and 7) were given in September, 1919. For purposes of comparison of two groups of different experience and maturity a test paper of the same total score as each high school paper was selected from the Alpha test blanks (Form 6) given to students in the University of Washington in March, 1919. The results of this comparison are shown in Table I.

TABLE I

	High School	University
Number of cases	168	168
Mean total score	105.6	105
Mean, Test 8	20.1	21.7
Mean age	15.7	21.5
Standard deviations		
of total score	26.0	26.0
of Test 8	9.9	7.2
of ages	1.6	5.1
Coefficients of correlation		
of total score and Test 8	.69 ± .027	.66 ± .029
of total score and age	.11 ± .05	−.16 ± .05
of age and Test 8	−.023 ± .05	−.06 + .05
Per cent Test 8 is of total score	19.0	20.5

These results show that with the groups used there was no undue influence of age or experience upon score in the information test. In both groups the correlation of age with total score and with information is, in effect, zero. In freedom from the influence of ages as such the information test suffers nothing as compared with the seven other tests of Alpha.

In the interpretation of these results it must be borne in mind that the university students whose papers were selected for pairing with the high school pupils were obviously an inferior group. Seventy-four per cent of the total university body exceeded the mean score of this group of 168 students. This gives a comparison between normal high school students and somewhat dull university students. The variabilities of the two groups are sufficiently close, except in the case of age, to make possible a direct comparison of coefficients of correlation. As all the correlations involving age are, in effect, zero the larger age variability of the college group may be ignored. No coefficients of partial correlation have been worked out for these data because of the absence of correlation between age and the other two variables.

In Table II are shown the scores made by pupils of different ages. By computing the ratio of total score to Test 8, and using year 14 as the norm, the correlation that must be added to total score in order to have the information test contribute the same relative amount for each age has been computed. As the table shows, the general tendency is for the information test to contribute a relatively large amount to the scores of the younger pupils, so that, while the difference is very slight, they appear to be somewhat favored by the information questions.

TABLE II

Age	Number of cases	Mean total score	Mean Test 8	Ratio	Correction
11–12	4	86.75	17.2	.1988	−1.1
13	7	105.28	21.7	.2062	−1.8
14	32	100.06	18.9	.1892	0
15	32	104.03	20.3	.1952	−0.7
16	33	111.51	20.0	.1801	1.0
17	34	107.08	19.6	.1837	0.6
18	18	108.88	21.1	.1998	−0.6
19	4	102.75	16.7	.1630	2.7

A negative correction indicates the amount by which total score is disproportionately large owing to the group doing relatively better than the 14-year group in Test 8.

To discover the effect of maturity on Test 8, Alpha, within a university group a selection of 736 papers was made from the same testing as in the preceding study. The ratio of total score and information score was computed as before, and also corrections prepared as before, using the 18-year group as the norm. The correction is the amount that must be added algebraically to the total score of its corresponding age to eliminate the influence of age upon information. It will be noticed from Table III that the correction for the age between 17 and 27 is less than 1 point of total score. For the oldest group, containing ages between 28 and 54, with the mean age 33, the correction is possibly large enough to be of significance.

To discover whether information is influenced by school grade in any unusual way a random selection from the college testing previously used was made. The results, which are tabulated in Table IV,

TABLE III

Age	Number of cases	Mean total score	Mean Test 8	Ratio	Correction
17	64	141.1	28.3	.2007	.43
18	111	132.8	27.0	.2043	0
19	111	136.0	28.5	.2099	−0.92
20	111	135.7	28.1	.2069	−0.30
21	111	129.3	27.2	.2109	−0.94
22	111	134.9	28.3	.2098	−0.91
23	111	132.6	27.4	.2045	−0.32
24	111	134.5	28.1	.2089	−0.90
25–27	63	136.6	28.8	.2106	−0.94
28–54	54	129.6	29.6	.2260	−1.96

show that in the case of each class the relative score contributed by Test 8 is approximately the same as for each other class.

TABLE IV

	Freshmen	Sophomores	Juniors	Seniors
Number of cases	193	165	163	120
Median age	19.5	20.6	21.7	23.3
Median total score	131.1	138.4	135.7	144.0
Median Test 8	28.8	30.0	29.8	30.1
Per cent Test 8 is of total score	21.1	21.6	21.9	20.9

Insofar as these high school and college test papers are typical, an analysis of the scores indicates that the information test is subject to the same influences and to the same degree as the other tests in Alpha. Neither chronological age nor school grade appears to influence unduly the score in the general information test except for a possible slight advantage of a group of unusually old and rather slow college students. With this exception, when the Alpha grades of students of different ages and school grades are compared the correction for the influence of age and experience is so small as to be negligible.

The high correlation of information tests with other measures of intelligence has made them valuable members of many batteries of tests. The ease with which an information test can be adjusted to a large number of needs by "trial and error" selection from a large number of questions makes it adaptable to many varied conditions. It is highly fortunate that the suspicion of these tests, based on common sense grounds, is unfounded in the case of group tests given to high-school and college students.

AN EXTENSIVE EXPERIMENT IN MOTOR LEARNING AND RE-LEARNING

S. R. BRADEN

University of Missouri

On April 25, 1921, an experiment in ball-tossing was begun and continued for 100 days. A box which stood 8 inches high with a 5-inch circular hole in the top served as the target. Balls of solid rubber $^{15}/_{16}$ of an inch in diameter were used. A cloth pad was placed in the bottom of the box to prevent the balls from bouncing out. The distance between the throwing line and the box was 12 feet. The experiment consisted simply in throwing the balls into the box. The ultimate object of the experiment was to study the nature of the learning capacity in trial and error learning when the learning is spread over a considerable period of time, and especially to note the course of the re-learning curve.

Twenty-five balls were used. When these had been thrown they were gathered and thrown again until 200 balls had been thrown. This constituted one day's practice. The score indicated only the balls which fell into the box. The experiment was begun on April 25, 1921 and ended August 25, 1921. Nine o'clock in the morning was the hour for daily practice. Sunday was taken as a rest day. The experiment lasted for 100 days, so that 20,000 balls were tossed. The scores were: 37, 25, 42, 45, 53, 51, 56, 55, 56, 60, 54, 65, 64, 65, 68, 69, 75, 79, 64, 70, 68, 67, 74, 68, 74, 76, 76, 76, 80, 79, 78, 81, 79, 82, 84, 85, 78, 84, 89, 85, 82, 88, 88, 87, 90, 88, 86, 86, 86, 89, 89, 90, 93, 92, 92, 90, 88, 88, 94, 96, 90, 91, 89, 92, 88, 88, 87, 90, 88, 86, 88, 87, 91, 89, 91, 89, 91, 91, 88, 88, 90, 89, 95, 93, 97, 98, 102, 98, 99, 102, 102, 102, 98, 91, 94, 89, 94, 94, 91, 95.

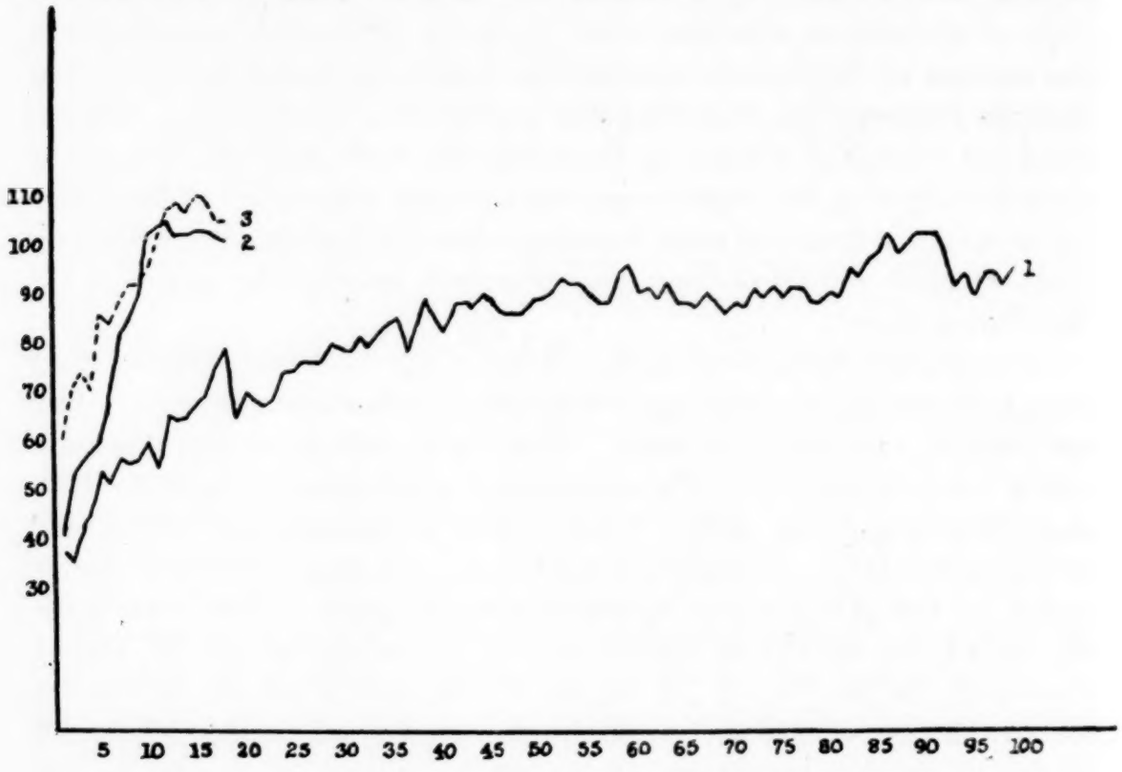
On March 5, 1923, 22 months and 11 days after the first experiment ended, a re-trial was carried on for 18 days under the same conditions as was the original experiment. The scores were: 40, 52, 56, 58, 60, 70, 81, 85, 90, 102, 104, 105, 102, 102, 103, 103, 102, 101.

Again on October 15, 1923, 6 months and 20 days later, a second re-trial was carried on for 18 days. The scores were: 60, 70, 74, 70, 86, 84, 89, 92, 92, 96, 102, 107, 109, 107, 110, 110, 105, 105.

The graph indicates clearly what the results were. The number of balls accurately thrown in the first 10 practices of the first experiment was 490; in the first re-trial it was 694; in the second re-trial it was 813. This gives an idea of the rapid rise of the curve in the re-trials. If we

reckon the total number of balls thrown for the first 18 practices of the first experiment we have 1029. The total for the first re-trial is 1516; and for the second re-trial is 1668. It is evident, therefore, that improvement was rapid in the retrials; and that the second re-trial shows a marked improvement over the first re-trial.

Many interesting questions arise. Can one gain much more than 50 per cent control in such a performance? Why, after one has thrown six balls successfully, does he miss the seventh? Is it better to use a



Graph 1 shows the daily record for 100 days' ball-tossing.
Graph 2 shows the daily record for 18 days' ball-tossing 22 months after the first experiment was finished.
Graph 3 shows the daily record for 18 days' ball-tossing 6 months and 20 days after the second experiment was finished.
The practices are represented on the horizontal axis and the scores on the vertical axis.

simple arm movement, or to bring the entire body into a rhythmic sway before the ball is released? Could one teach others to toss the balls accurately; or must each learner, by the trial and error method, discover his own "best way?"

The experiment has to do with the vital elements of character. A high degree of concentration is required and it is easy to "let down."

Perhaps more fruitful results would be obtained if all good throws counted according to a graduated scheme, as in target-shooting. This would tend to do away with discouragement, especially if the experiment were used for children. For the strictest of accuracy, only one ball should be used, which could be returned to the thrower by means of a trough. If balls of absolute identity in weight and size can be secured, of course they would be preferable to the one ball with the trough. The experiment can be made as difficult or as easy as is desired by changing the throwing distance. Since only inexpensive apparatus is required, it is hoped that many others will perform this experiment. It can be used either in home or school and throws much light upon improvement and re-learning.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY CECILE COLLOTON

Department of Educational Psychology, The Lincoln School of Teachers College

Intelligence Tests

13:77 *versus* 12:05; *a Study in Probable Error.* Maurice B. Hexter and Abraham Myerson. Mental Hygiene, 1924, January, 69–82. A critical discussion of Brigham's "A Study of American Intelligence" from the point of view of race discrimination and propaganda.

The Army Tests and the Pro-Nordic Propaganda. Wm. C. Bagley. Educational Review, 1924, April, 179–187. Challenges Brigham's interpretations and assumptions based on army test data in his book "A Study of American Intelligence." Proposes in contrast, "the constructive program of the rational equalitarian."

The History of Mental Testing. Kimball Young. The Pedagogical Seminary, 1924, March, 1–48. Summarizes the history of psychology in Germany, France, England, and America with brief mention of the names, dates, and work of the men whose contributions are outstanding.

The Mental Test as a Psychological Method. Lewis M. Terman. Psychological Review, 1924, March, 93–118. Discusses the value of the mental test in relation to the whole science of psychology.

A Psychological Description of Intelligence. Raymond H. Wheeler. Psychological Review, 1924, March, 161–174. Defines intelligence as "derived covert behavior" on four levels of complexity; localization, recognition, comprehension, and interpretation.

Intelligence Tests versus Entrance Examinations as a Means of Predicting Success in College. Ada Hart Arlitt and Margaret Hall. The Journal of Applied Psychology, 1923, December, 330–338. A three years' study of scores on entrance examinations, college grades, and scores on intelligence tests of 305 college students shows entrance grades to be the best measure on which to predict success in college.

Educational Tests

The New Type Examinations in Physics. H. W. Farwell. School and Society, 1924, March 15, 315–322. Illustrates the use of true-false, completion, and multiple choice tests in physics as satisfactory supplements to the old type of examination.

The Relation between Reading Ability as Measured by Certain Standard Tests and the Ability Required in the Interpretation of Printed Matter Involving Reason. B. T. Ritter and W. T. Lofland. The Elementary School Journal, 1924, March,

529–546. Statistical data support the contention that there are different types of reading ability requiring specialized types of training.

The Reliability of Measurements in Spelling. R. L. Morton. The Journal of Education Method, 1924, April, 321–328. Questions the grade standards and the assigned difficulties of words in the Ayres Scale. Data secured from 2863 children in cities in Ohio.

Fundamental Issues in Evaluating Composition. Joseph C. Gainsburg. The Pedagogical Seminary, 1924, March, 55–77. Points out the inadequacies and errors in composition scales and suggests a list of criteria for judging a composition.

The Reliability of Tests Requiring Alternative Responses. Wm. Asker. Journal of Educational Research, 1924, March, 234–240. The present method of scoring—rights minus wrongs—does not eliminate the effect of guessing in individual cases. Guessing should be discouraged.

Case Studies

Reginald. Ross D. Murphy. The Psychological Clinic, 1924, January, 154–156. Reports the study of a twelve-year-old boy who is "nervously a wrecked machine which refuses to function."

James, A Study of a Post-encephalitic Child. Rose Wimenitz. The Psychological Clinic, 1924, January, 144–148. Details of the study with the prognosis a question.

The Spotless Child. Beatrice Green. The Psychological Clinic, 1924, January, 149–153. The story of Marie, a middle grade imbecile whose mental condition is explained in part by a long series of illnesses very early in her life.

What Are Little Boys Made of? Ellen M. Jordan. The Psychological Clinic, 1924, January, 138–143. A case study of reading disability.

The Superior Child. Alice M. Jones. The Psychological Clinic, 1924, January, 130–137. Reports four case studies of children who have general intellectual superiority and marked musical ability.

A Case of Secondary Mental Deficiency with Musical Talent. Blanche M. Minogue. The Journal of Applied Psychology, 1923, December, 349–352. Describes a male patient at Letchworth Village—chronological age 23, mental age about 7 years—who is a pianist of unusual ability.

The Problem Child. Phyllis Blanchard and Richard H. Paynter, Jr. Mental Hygiene, 1924, January, 26–54. Reports the results of a study of 500 school children considered problems at home or at school. Medical, psychiatric, and psychological examinations were given and a social history of each child obtained. Seven case studies are reported in detail.

Problem of the Quasi-delinquent in the School. Dorothy Wallace. Mental Hygiene, 1924, January, 115–165. What schools can do and are doing to prevent delinquency. Illustrated by case studies.

New Tests

Rhode Island Intelligence Test. G. E. Bird and C. E. Craig. Group test for children from three to six years of age. Public School Publishing Company. Bloomington, Illinois.

New Standardized Oral Reading Check Tests. William S. Gray. Supplied in pairs of 100 tests for testing 20 individuals five times. Set 1, Grades I and II.

Set II, Grades II–IV. Set III, Grades IV–V. Set IV, Grades VI–VIII. Price of each set including direction sheets and record sheets $1.50. Public school Publishing Company, Bloomington, Illinois.

Stevenson Problem Analysis. P. R. Stevenson. For training in analysis of word problems in arithmetic. Tentative norms. Public School Publishing Company, Bloomington, Illinois.

Pressey Test in Diagnostic Composition. S. L. Pressey, Mrs. E. U. Bowers, Helen Ruhlen and F. R. Conkling. Separate tests for Capitalization, Punctuation, Grammar, and Sentence structure. Public School Publishing Company, Bloomington, Illinois.

Iowa Physics Test. Dr. Harold L. Camp. Heat, Mechanics, Electricity, and Magnetism. Two forms for each test. Tentative norms. Public School Publishing Company, Bloomington, Illinois.

Chemistry Test Epsilon. S. G. Rich. Second Form. Twenty-six items with multiple choice answer. Public School Publishing Company, Bloomington, Illinois.

Miscellaneous

The Relation of Mental to Physical Growth. William Filler Lutz. The Psychological Clinic, 1924, January, 125–129. Reports the results of mental and physical examinations of 100 boys.

The Conservation of Talent. Lewis M. Terman. School and Society, 1924, March 29, 359–364. A summary of the results to date of an investigation of 1000 superior children. Objective data along anthropometric, medical, educational, social and psychological lines are being collected and the histories of the children are to be followed into adult life as far as possible.

Diagnostic and Cumulative Surveys of School Achievement in Minnesota. M. J. Van Wagenen. 1924, March, 147–168. Points out and illustrates significant features of school surveys which are of importance to grade principals and school superintendents.

Sex Differences and Sex Variability in Learning Capacity. W. H. Pyle. School and Society, 1924, March 22, 352. Studies of boys' and girls' learning capacity in card sorting and digit-letter substitution show the girls' scores to be consistently higher. Learning of this type improves with age.

Selection in "Trial and Error" Learning. William R. Wilson. Psychological Review, 1924, March, 150–160. Explains the origin and principles of selection of movements in trial and error learning in terms of familiar processes.

The Franz Dot Tapping Test as a Measure of Attention. Irene Case Sherman. The Journal of Applied Psychology, 1923, December, 353–359. Discusses the use of the Franz Dot Tapping Test with 172 unselected cases. Results are compared with chronological age, mental age, and IQ. Six tables present statistical data.

Organizing Classes for Gifted Children. Herbert D. Bixby. Educational Review, 1924, April, 196–199. Discusses the selection of gifted children for special classes, the importance of a superior type of teacher, and the ways in which the public can aid the work.

Prevention of the Misfits. Margaret S. McGill. The Pedagogical Seminary, 1924, March, 49–54. Emphasizes the need of practical educational advantages and competent vocational guidance for foreigners.

A Study of Kraepelins' Continuous Substraction Test. The Journal of Abnormal Psychology and Social Psychology, 1924, January–March, 385–388. Trial with 113 unselected children show this test to be applicable as an attention test only for the normal adult.

Two Hundred Cases of Dementia Praecox Tested by the Stanford Revision. Mary M. Wentworth. The Journal of Abnormal Psychology and Social Psychology, 1924, January–March, 378–384. Results of tests with dementia praecox patients indicate volitional and affective disorders rather than intellectual defects and emphasize the importance of individual study and treatment at an early stage of the disease.

The Influence of Latin on the Spelling of English Words. Warren W. Coxe. Journal of Educational Research, 1924, March, 223–233. An experiment conducted with five groups of classes each following a different procedure shows that the study of Latin has an appreciable effect on English spelling, especially when rules governing the spelling of Latin derivatives are developed.

Reducing the Variability in Teachers' Marks. E. J. Ashbaugh. Journal of Educational Research, 1924, March, 185–198. Reports an experiment designed to show student teachers the great variability in teachers' grading. Suggests clear-cut definition of what is being marked and agreement as to the relative value to be assigned to the various elements as a means of improvement.

The Memory Value of Charts in College Lecture Courses. Donald A. Laird. School and Society, 1924, March 29, 381–382. The use of charts with a lecture, unless directly referred to and elaborated upon, is not effective.

Some Whys of Whims—A Genetic Study. Garry Cleveland Myers. The Pedagogical Seminary, 1924, March, 78–83. Notes on the behavior of a child showing that whims are apparently just habits.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY LAURA ZIRBES

Educational Psychology Simplified

The Child's Mind and the Common Branches, by Daniel Wolford La Rue. New York: Macmillan, 1924. Pp. X + 483.

"*Practical and positively helpful,* is the motto I have placed before myself in writing this book," announces Dr. La Rue in the opening sentence of his introduction. For once a motto has been effective, and the result is an elementary textbook in educational psychology that teachers in training at normal schools should be able to read with comprehension and later to consult with hopefulness.

Part I (118 pages) sketches the physiological basis of mental activity, develops the concept of bonds, and discusses the three laws of Readiness, Exercise and Effect as formulated by Professor Thorndike. The remainder of the book makes application of this matter to the teaching of the common branches. Inquiry is made as to the desirable bonds to be formed in each subject, the proper order of their formation, and the practical significance of the laws of learning in determining method. The ideas presented are, for the most part, in accord with the findings of modern research, and are traceable to their sources in excellent reference lists at the ends of chapters. In the case of reading alone, where oral training is stressed, is there any marked difference in emphasis from that of recent expert treatments.

An engagingly personal style and diction, together with an abundance of homely, pointed illustration, is calculated to hold the reader's interest and quicken his understanding. Many supplementary exercises are suggested to illuminate and fix more securely the principles enunciated in the text. There are, of course, obvious limitations to the value of such a book for the mature student of education; but for the beginner, the young high school graduate just entering the normal school, its spirit and method seem especially well conceived.

M. H. Willing.

A Teaching Help for Educational Psychology

Syllabus for a First Course in Educational Psychology, by F. B. Knight and G. M. Ruch. Iowa City, Iowa: The Iowa Supply Co., 1924.

Teachers of educational psychology who use Gates' "Psychology for Students of Education" will welcome a recent syllabus designed to accompany this text. The authors have divided the material of the text into 30 lessons. For each lesson the assignment has been given in detail accompanied by notes explaining or calling attention to difficult points. Detailed discussions of the material for which the student should be held accountable are included in each lesson. The notes on the lesson show evidence of much thought and a knowledge of the difficulties with which students in beginning educational psychology meet. The points upon which the pupils are expected to recite are well selected.

The only apparent weakness in the syllabus is the lack of references to supplementary reading material.

Carl L. Huffaker.

University of Arizona.

The Effect of Attitude on Women's Physical Freedom

Woman's Physical Freedom, by Clelia D. Mosher. New York: The Woman's Press, 1923. Pp. 87.

In its refutation of the fallacious psychology of sex limitations which has been disseminated among women by tradition and precept, Dr. Mosher's admirably sane and scientifically sound treatment of women's disabilities should prove a valuable contribution to the psychology of individual differences.

The book is written from the medical standpoint and Dr. Mosher bases her conclusions on the study of more than 2000 university women, her observations having extended over a period of 30 years. Her thesis is, in brief, that the physical freedom of the woman of the present college generation "with her perfectly working body and superb health" is a direct outcome of "an alteration of the morbid attitude" of women themselves toward factors which produce disabilities.

The assumption that certain sex differences are due to original nature has influenced largely our educational theories and practices.

The author has shown that even these physical traits which have been thought to cause differences in behavior, achievement and mental activities, are themselves largely determined (functionally) by ideas. Attitudes imposed by tradition and fostered by training account for some differences which have been considered fundamental.

Emancipation from the deleterious effects of regarding functional periodicity in the light of a physical disability—as though digestive disturbances were to be regarded as normal and inevitable—the increase in height (the average height of college women today is 1.2 inches above the average height 30 years ago) and size of waist, resulting from freedom from earlier constrictions of dress and the greater physical vigor, together with the increased muscular activity of the modern athletic woman constitute important factors in woman's physical freedom which in a "world of unlimited opportunities" will enable "the woman of the rising generation to answer the question of what woman's real capacities are."

It is significant that Dr. Mosher, in pointing out that it rests with the woman alone whether she rejects this new physical freedom, clinging to the old ideal of weakness and dependence, places first "the opportunities of making herself a better citizen," that is, of using "the opportunities for her self-development—interesting and worthwhile living, service to the race, economic and civic—in other words, to have a life crowded full of things worth having."

The book is specially commended to students and teachers of educational psychology. It presents a point of view fundamentally sound and significant in the investigation of the direct share of sex in the production of mental differences.

MAUD A. MERRILL.

TYPICAL LABORATORY EXPERIMENTS IN THE PSYCHOLOGY OF LEARNING

A Laboratory Manual in the Psychology of Learning, by William Henry Pyle. Baltimore: Warwick and York, 1923. Pp. 161.

"This manual is published in the hope of furthering and facilitating experiments in educational psychology and is the outcome of 14 years of experimentation."

In the first chapter the nature, aims and method of experimentation are presented in a clear, simple and forceful way which no one seems to have excelled. The first few pages are particularly good.

Thirteen pages are given to "Mathematical Treatment of Data" covering briefly the learning curve, frequency surface, measures of central tendency, measures of variability, and correlations, with a small amount of illustrative materials. In the next 120 pages the author lists experiments in a rather elaborate and effective manner, stating the object, materials, method and the results. Ball-tossing and mirror-drawing are given as trial and error types of motor learning followed by a card sorting and a marble sorting experiment. Then come the familiar digit-symbol experiment to illustrate semi-motor learning. Next the card sorting experiment is adopted to measure "retention of motor learning" and to "study inhibition."

Following these are experiments in which the tachistoscope is used. The spot pattern test is used to measure perception span. Serial learning is studied with non-sense syllables and meaningful words. In an experiment on associative learning, paired associates are used. Visual and auditory methods of rate and ideational learning are compared.

If the purpose is to master some technique, manipulate data and laboratory apparatus apart from any definite bearing upon everyday problems of life and learning this is a pretty good manual.

It seems a pity however, that the author does so little to prevent the innocent student from thinking that all learning curves are of the concave type or that Professor William Henry Pyle has had to struggle alone for these 14 years to create out of nothing this group of rather ingenious experiments. Nevertheless, for models of simple statement of technique, any one who desires to write up a series of laboratory experiments in psychology will do well to turn to this little manual by Professor Pyle.

GARRY CLEVELAND MYERS.

Cleveland School of Education.

YERKES-BRIDGES POINT SCALE REVISED

1923 *Revision: A Point Scale for Measuring Mental Ability* by Robert M. Yerkes and Josephine C. Foster. Baltimore: Warwick and York, 1923. Pp. VII + 219.

The original edition of this book described in a series of papers the point scale method, the use of the original Pre-Adolescent Scale, and the revision based on an analysis of the results. In the final chapter of that volume the authors proposed to develop a universal scale based

on three perfected limited-range scales—the Pre-Adolescent, the Adolescent-Adult, and the Infant or Child Scale. The present volume presents these three scales. The Pre-Adolescent scale is essentially as in the first edition and the two supplementary scales are submitted merely in tentative form.

The material has been entirely reorganized and rewritten. The chapters of the first edition which dealt with the history of the method, the initial form of the scale, and the results upon which the Pre-Adolescent Scale was based are omitted. A brief description of the Point Scale Method and a comparison with the Binet Scales is followed by directions for giving the Pre-Adolescent Scale. In this chapter the rewriting, which follows closely Terman's style in "The Measurement of Intelligence," has been particularly effective. The standardized directions stand out clearly in italics and additional directions make clear to the examiner exactly what is to be done. Many illustrative responses have been added and these, together with the detailed "remarks" on practically every test, aid the examiner materially in the interpretation of a performance. The results of the original scale are summarized in one chapter and norms based on these results are recommended for use as standards of comparison. They are repeated at the end of the book for ready reference, together with the printed materials necessary in the giving of the test. The Adolescent-Adult and the Infant Scales are described and full directions for their use are given. Examiners who feel the need of more adequate measures of the years above twelve or thirteen or below seven will be interested in using these scales critically and in aiding in their revision.

An especially interesting part of the book is the section given to a discussion of "The Binet versus the Point Scale Method of Measuring Intelligence." A careful analysis of the two scales shows how radically they differ on fundamentally important points, and points out the advantages in each scale which might help to make a "new scale infinitely superior to anything Binet or his successors have imagined."

CECIL M. COLLOTON.

A TIMELY SURVEY OF THE FIELD OF HEARING

Hearing by Robert M. Ogden. New York: Harcourt, Brace and Co., 1924. Pp. XIII + 351.

This volume has been many years in the making, and is the work of an authority whose intimate association with the work in this field, as

editor in charge of the department of the Psychological Bulletin dealing with the sense of hearing, eminently qualifies him for such an undertaking.

It has the unique distinction of being the only work of its kind in any language, representing a systematic interpretation and organization of the increasingly large number of experimental studies in this field, as well as that of being a lucid and trustworthy account of its many ramifications. The book comprises a discussion of the physics and physiology of hearing, an analysis of tone, vocal sound and noise, and the parts they play in music and language, localization of sound, pathology of hearing and aural education. Notwithstanding the range of treatment, which ordinarily would lead to superficiality and arbitrariness, one is favorably impressed with the author's attempt to marshall the significant facts and to maintain an attitude of open-mindedness on all debatable issues. The extensive list of sources, utilized and specifically referred to, is an outstanding feature.

One must appreciate the fact that the author has kept the varied interests of a wide range of readers constantly in mind. Consequently, this contribution should prove an indispensable aid, not only to the psychologist and otologist, but to the teacher and student of music and phonetics in the solution of his special problems as well.

J. KUDERNA.

CURRICULUM-MAKING FOR CHINA

Education for Modernization in China, by Peng Chun Chang. New York: Teachers College, Columbia University, 1923. Pp. 92.

The task to which the author of this book addresses himself is the formulation, through analysis and constructive types of reasoning, of a plan for curriculum revision in China's middle schools. China has great need of leaders in her effort to become modern. The present tendency toward undiscriminating imitation of the forms and materials of Western secondary education is not likely to meet this need successfully. To be safe, it is necessary first to answer such qustions as these: What are the significant characteristics of this modernism that China so much desires? What influences produced them? How can these influences, probably adventitious in the West, be consciously generated in Chinese schools so as to develop in prospective leaders the essential qualities of the modern spirit?

The author goes to the foremost analysts of the day for an interpretation of modernism. He concludes from this research that scientific method, individuality, and democracy (the three chief characteristics of Western modernism) have been the direct outgrowths of the "exploring, pioneering, frontier community-building, large-scale producing, and the like" incident to *Expansion* in the last four or five centuries. On the basis of this hypothesis he next sets up criteria for the selection of curriculum materials (activities) that will serve in a Chinese environment as substitutes for the influences identified in the expanding life of the Western nations, particularly of America. Finally he reaches the point of offering suggestive outlines of projects, formulated subject matter, and types of social organization among pupils. These outlines are purely tentative and intended only for the initial direction of experimentation.

The study as a whole is a remarkably fine illustration of constructive thinking, very clear in its progressions and not too far removed from practical needs and possibilities. It would be hard to find a more thorough or consistent application of the educational philosophy of Professor Dewey and the Columbia group to the exigent problem of curriculum revision not only in China but in America as well.

M. H. WILLING.

A HANDBOOK ON READING METHODS

How to Teach Reading, by Mary E. Pennell and Alice M. Cusack. New York: Houghton Mifflin, 1923.

Based on the theory that, for ordinary purposes, a teacher needs to know the results and not the processes of psychological research, Pennell and Cusack's new book on reading bridges the gap between the laboratory and the classroom. The conclusions of the psychologists are stated briefly and without technical language. The results of investigations concerning the work of the eye in reading are admirably summed up. The authors consider the laws of learning only as they apply to reading. Throughout, the book is based upon sound psychological principles, which it accepts as a major premise and from which it proceeds deductivity. *How to Teach Reading* is distinctly the product of persons who understand the problems of the classroom. A chapter

on "Reading Stimuli" makes specific and usable suggestions about greetings, illustrations, puzzles, chart stories, etc. Various methods of keeping grade records, class and individual graphs are illustrated by simple diagrams which any energetic teacher could use with a minimum of effort.

In organization, the book is admirably suited to reference needs. As far as possible, outlines and summaries are used, so spaced that the eye quickly locates what is wanted. The waste of repetition is avoided by careful cross references which make all related material easily available.

The writers have great confidence in the power of well framed "motive" questions to bring about a state of readiness for reading They also emphasize the functions of questions which demand judgment and organization. The use of these three types of questions in carefully selected reading units for the elementary grades is well illustrated throughout the book.

One of the great contributions of recent studies in the psychology of reading has been to show that reading is a complex of many specific abilities. General training or specific training along one line will not result in adequate control of all of these abilities but will facilitate in some cases and inhibit in others. This fact implies that careful analyses must be made. (1) Of the specific abilities involved; (2) the ways and means by which they develop, (3) and the peculiar reactions which result from over or under training. From this standpoint the material on diagnosis and remedial work is inadequate except for very elementary work. The section on the between recitation period will also be disappointing to many because it contains little new material and avoids any suggestions regarding the primary difficulty involved in children's undirected work, namely that of providing self-checking devices for material of intrinsic value to the child.

Because of the directness and simplicity with which the psychological principles are stated, and because of the wealth of illustrative material used, the book will give immediate and practical help to many teachers and will no doubt be used widely and successfully in normal schools as a starting point for discussions of this most complex problem, "How to Teach Reading."

RUTH I. BROWN,
ELIZA OGLESBY.

Teachers College, New York City.

Other Publications Received

Coe, George A. *Law and Freedom in the School.* Chicago: The University of Chicago Press, 1924, pp. 133.

Coy, Genevieve Lenore. *The Interests, Abilities and Achievements of a Special Class for Gifted Children.* New York: Teachers College, Columbia University, 1923, pp. 194.

Cunningham, Bess V. *The Prognostic Value of a Primary Group Test.* New York: Teachers College, Columbia University, 1923, pp. 54.

Dresser, Horatio W. *Psychology—In Theory and Application.* New York: Thomas Y. Crowell Co., 1924, pp. 727.

Foster, William S. *Experiments in Psychology.* New York: Henry Holt, 1923, pp. 309.

Handschin, Charles H. *Methods of Teaching Modern Languages.* Yonkers, N. Y.: World Book Co., 1923, pp. 479.

Hughes, Elizabeth, *Children of Preschool Age in Gary, Indiana.* Washington Government Printing Office, 1922, pp. 175.

Jonckeere, Tobie. *La Pedagogie experimentale.* Bruxelles: Maurice Lamerton, 1924, pp. 168.

Pierce, Frederick. *Mobilizing the Mid-brain.* New York: Dutton, 1924, pp. 259.

Prescott, Daniel Alfred. *The Determination of Anatomical Age in School Children and Its Relation to Mental Development.* Cambridge, Mass: Published by The Graduate School of Education Harvard University, 1923.

Rignano, Eugenio. *The Psychology of Reasoning.* New York: Harcourt, Brace and Co., 1923, pp. 395.

Somers, Grover Thomas. *Pedagogical Prognosis.* New York: Teachers College, Columbia University, 1923, pp. 129.

Thurstone, L. L. *Public Personnel Studies.* Washington, D. C.: Bureau of Public Administration Institute for Government Research, $3.00 per year, Vol. I, Nos. 1, 4, 5, 6.

Toops, Herbert A. *Tests for Vocational Guidance of Children Thirteen to Sixteen.* New York: Teachers College, Columbia University, 1923, pp. 159.